IJCSIS Vol. 12 No. 2, February 2014
ISSN 1947-5500

International Journal of Computer Science & Information Security


Cornell University
Library


Cogprints


Google scholar


.docstoc
find and share professional documents


ScientificCommons

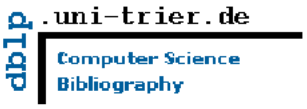

View my documents on
Scribd


BASE
Bielefeld Academic Search Engine


SCIRUS
search engine for science


SciRate.com


CiteSeerx beta


dblp .uni-trier.de
Computer Science
Bibliography


Q·Sensei


DOAJ
DIRECTORY OF
OPEN ACCESS
JOURNALS


EBSCO
HOST


ProQuest

IJCSIS

ISSN (online): 1947-5500

Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results.

CALL FOR PAPERS
International Journal of Computer Science and Information Security (IJCSIS)
January-December 2014 Issues

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.
See authors guide for manuscript preparation and submission guidelines.

Indexed by Google Scholar, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Scopus Database, Cornell University Library, ScientificCommons, ProQuest, EBSCO and more.

Deadline: see web site
Notification: see web site
Revision: see web site
Publication: see web site

Context-aware systems
Networking technologies
Security in network, systems, and applications
Evolutionary computation
Industrial systems
Evolutionary computation
Autonomic and autonomous systems
Bio-technologies
Knowledge data systems
Mobile and distance education
Intelligent techniques, logics and systems
Knowledge processing
Information technologies
Internet and web technologies
Digital information processing
Cognitive science and knowledge
Agent-based systems
Mobility and multimedia systems
Systems performance
Networking and telecommunications
Software development and deployment
Knowledge virtualization
Systems and networks on the chip
Knowledge for global defense
Information Systems [IS]
IPv6 Today - Technology and deployment
Modeling
Software Engineering
Optimization
Complexity
Natural Language Processing
Speech Synthesis
Data Mining

For more topics, please see web site https://sites.google.com/site/ijcsis/

For more information, please visit the journal website (https://sites.google.com/site/ijcsis/)

Editorial
Message from Managing Editor

The ***International Journal of Computer Science and Information Security*** *(IJCSIS), is an international forum for scientists and engineers involved in all aspects of computer science and technology to publish high quality, refereed papers. As a scholarly open access peer-reviewed international journal, IJCSIS encourages emerging scholars and academicians globally to share their professional knowledge, innovative ideas and research in the fields of computer science, engineering, technology and related disciplines. The objective of IJCSIS is to bridge the research community and technology developers from academia and industry through submitting/publishing their research–based papers, articles and case reviews on various topics of current concern on Computer Science, Security and Technology.*

IJCSIS archives all publications in major academic/scientific databases; abstracting/indexing, editorial board and other important information are available online on homepage. Indexed by the following International agencies and institutions: Google Scholar, Bielefeld Academic Search Engine (BASE), CiteSeerX, SCIRUS, Cornell's University Library EI, Scopus, DBLP, DOI, ProQuest, EBSCO. Google Scholar reported increased in number cited papers published in IJCSIS. IJCSIS supports the Open Access policy of distribution of published manuscripts, ensuring "free availability on the public Internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of [published] articles".

IJCSIS editorial board ensures a rigorous peer-reviewing process and consisting of international experts. We value your support immensely and invite you to be a part of IJCSIS research movement.

We look forward to your collaboration. For further questions please do not hesitate to contact us at ***ijcsiseditor@gmail.com****.*

A complete list of journals can be found at:
http://sites.google.com/site/ijcsis/

IJCSIS Vol. 12, No. 2, February 2014 Edition

ISSN 1947-5500 © IJCSIS, USA.

Journal Indexed by (among others):

IJCSIS EDITORIAL BOARD

TABLE OF CONTENTS

device. These are, multiplier using carry save adder (CSA) and vedic multiplier. The proposed Dual field ECC processor using Vedic multiplier can reach a speed up to 151MHz.

Keywords - Dual Field ECC; Galois Field; Elliptic curve scalar multiplication; Elliptic curve point addition; Elliptic curve point doubling; Dual field multiplier; Dual field adder; Vedic multiplier; FPGA.

4. Paper 31011422: Wavelet based Iris Recognition for Biometric Systems (pp. 25-29)

Jitendra Singh Uriaya, Department of Electronics and Communication, Jabalpur Engineering College, Jabalpur India (482011)
Agya Mishra, Department of Electronics and Communication, Jabalpur Engineering College, Jabalpur India (482011)

Abstract ─ Iris recognition is a critical area of research in the field of security system and personal identification. In this paper, a novel, efficient technique for iris recognition is presented. The aim is to develop a lifting wavelet based algorithm. This method reduces the noise to the maximum extent possible, and extracts important information from the image. In this paper recognition is being done biorthogonal wavelet based method which is then compared with other wavelet family. The method was tested on the CASIA dataset of iris image. For matching the iris template Hamming Distance technique is used and checks the accuracy of the system. The present algorithm provides the 94.4% recognition rate and accuracy approx. 99 % on the CASIA iris data base version 1.

Keywords─ Iris recognition, Biorthogonal wavelet, Haar wavelet, Daubechies wavelet, Iris template, Hamming distance, Match score.

5. Paper 31011429: Performance of Watermarking System using Wavelet Column Transform under Various Attacks (pp. 30-35)

Dr. H. B. Kekre, Senior Professor, Computer Engineering Department MPSTME, NMIMS University, Vile Parle, Mumbai, India,
Dr. Tanuja Sarode, Associate Professor, Computer Department, Thadomal Shahani Engg. College, Bandra, Mumbai 50, India
Shachi Natu, Ph.D. Research Scholar, Computer Engineering Department MPSTME, NMIMS University, Vile Parle, Mumbai, India

Abstract — In this paper a novel wavelet transform based color image watermarking technique has been proposed. Column wise DCT wavelet transform has been applied to cover image. A watermark of size 128*128 is compressed and embedded in the cover image of size 256*256 using DCT wavelet column transform. Compression of watermark reduces the payload of information to be embedded in the cover image without significant loss of information. Robustness of proposed technique is tested against three types of attacks namely, cropping, Binary distributed run length noise and Gaussian distributed run length noise. Performance of proposed technique is observed for various compression ratios 1.45, 1.6, 1.77, 2, 2.285 and 2.67 beyond which the distortion in compressed watermark is perceptible. This technique proves to be exceptionally robust for Binary distributed run length noise attack. For Gaussian distributed run length noise attack also acceptable robustness is achieved. For cropping also it gives good performance except compression ratio 2. Higher compression of watermark results in better imperceptibility of watermarked image. Also higher compression ratio leads to increased robustness against attacks on watermarked image.

Index Terms — Watermarking, DCT wavelet, Column transform, Binary distributed Run length noise, Gaussian distributed Run length noise.

6. Paper 31011433: Content Based Image Retrieval Through Distance Similarity Metrics (pp. 36-40)

Praveen Mishra, Department of Computer Science, VNS Institute of Technology Bhopal, India
Prof. G.S Prajapati, Department of Computer Science, VNS Institute of Technology

Abstract - Searching Test Image from Image databases using features extraction from the content is currently an active research area. In this work we present novel feature extraction approaches for content-based image retrieval when the query image is color image. To facilitate robust man-machine interfaces, we accept query images with color attributes. Special attention is given to the similarity measure with different distance matrices properties since the Test Image and Object Image from database finding the distance measuring. Several applicable techniques within the literature are studied for these conditions. The goal is to present the user with a subset of images that are more similar to the Object Image. One of the most important aspects of the proposed methods is that the accuracy measurement of the different database images. This significantly improves the feature extraction process and enables the methods to be used for other computer vision applications.

Keyword - Color image, Image retrieval, Euclidean metrics, Manhattan metrics and Correlation.

7. Paper 31121329: Evolutionary Neural Network Model For Dynamic Channel Allocation In Mobile Communication Network (pp. 41-55)

Ojesanmi Olusegun and Ugege Peter, Department of Computer Science, Federal University of Agriculture, Abeokuta, Nigeria.

Abstract - Mobile communication aims at transmitting multimedia data and at the same time provides guaranteed quality of service (QoS) to all the applications. The challenge is to develop an efficient allocation scheme for assigning resources without compromising the QoS. In meeting this challenge, this paper proposes an evolutionary neural network approach with dynamic allocation to utilize frequency spectrum efficiently and to reduce call blocking probabilities.

Keywords: Neural Network, Mobile network, call blocking probabilities, dynamic channel, QoS, Multimedia.

8. Paper ID45 : Segmentation of Kannada Handwritten Text Line through Computation of Variance (pp. 56-60)

Sunanda Dixit #1, Suresh Hosahalli Narayan #2
1Information Science and Engineering Department, Dayananda Sagar, College of Engineering, Bangalore, India
2 Department of Instrumentation Technology, Bangalore Institute of Technology, Bangalore, India

Abstract — Handwritten text line segmentation is an important task of Optical Character Recognition. The proposal discusses a novel technique for Segmentation of Lines for the handwritten text document written in Kannada Language. The algorithm imbibes the approach of finding the Components, bounding box and computing the coefficient of variance. Preprocessing steps such as correction of skewness due to improper scanning as well as removal of noise is assumed to have been performed on the data. Our technique of line segmentation addresses the key complexity issue of variations in the gaps between and across lines as well as curly and skew character text of handwritten document. The purpose of this paper is to segment line by connecting the centroids within the bounding box and employing the coefficient of variance. The method has been tested on unconstrained Kannada scripts. An experimental result obtained by executing the proposed method of line segmentation has achieved high degree of accuracy and high performance.

Keywords— Centroids, Coefficient of Variance, *Document Image Processing, Optical Character Recognition, Text line segmentation*

9. Paper ID64: Incremental Classification using Feature Tree (pp. 61-65)

*N N Vadnere #, R G Mehta *, D P Rana *, N J Mistry *, M M Raghuwanshi ***
 # M.Tech, Computer Science Department, Manipal University Jaipur, Jaipur, India
** SVNIT, Surat, India*
***RGCET, Nagpur, India*

Abstract— In recent years, stream data have become an immensely growing area of research for the database, computer science and data mining communities. Stream data is an ordered sequence of instances. In many applications of data stream mining data can be read only once or a small number of times using limited computing and storage capabilities. Some of the issues occurred in classifying stream data that have significant impact in algorithm development are size of database, online streaming, high dimensionality and concept drift. The concept drift occurs when the properties of the historical data and target variable change over time abruptly in such a case that the predictions will become inaccurate as time passes. In this paper the framework of incremental classification is proposed to solve the issues for the classification of stream data. The Trie structure based incremental feature tree, Trie structure based incremental FP (Frequent Pattern) growth tree and tree based incremental classification algorithm are introduced in the proposed framework.

Keywords— Stream Data, Trie, Discretization, Incremental Classification, Feature-Tree

10. Paper ID69: Discretization of Temporal Data: A Survey (pp. 66-69)

*P. Chaudhari *, D. P. Rana *, R. G. Mehta *, N. J. Mistry *, M. M. Raghuwansh i#*
**SVNIT, Surat, India*
#RGCET, Nagpur, India

Abstract — In real world, the huge amount of temporal data is to be processed in many application areas such as scientific, financial, network monitoring, sensor data analysis. Data mining techniques are primarily oriented to handle discrete features. In the case of temporal data the time plays an important role on the characteristics of data. To consider this effect, the data discretization techniques have to consider the time while processing to resolve the issue by finding the intervals of data which are more concise and precise with respect to time. Here, this research is reviewing different data discretization techniques used in temporal data applications according to the inclusion or exclusion of: class label, temporal order of the data and handling of stream data to open the research direction for temporal data discretization to improve the performance of data mining technique.

Keywords— Temporal data, Discretization, Supervised, Incremental, Nonparametric

Security in VoIP

Elsayed M. Elshamy, Osama S. Faragalla, Sayed El-Rabaie, Osama Elshakankiry and Fathi Sayed

***Abstract*—**In recent years, result of the significant in the field of information technology, secure communication techniques have increased widely and unexpectedly, Specially with VoIP communication network. In order to establish reliable communication technology and to ensure that voice reaches its intended end and to be accessible to all through the shared network, there is a need to secure communication in VoIP communication networks and encrypt the transmitted voice "VOICE MAIL". The proposed system is performed in two stages, the first stage includes cryptosystem with Chaotic Encryption of voice. The second stage using fingerprint authentication.

***Keywords*—**VoIP, Fingerprint Authentication, Chaotic Encryption, Cryptosystem.

I. INTRODUCTION

FOR secure communication in the VoIP network as shown in Fig. 1, authentication and security of voice must be provided. In this paper, we use a cryptosystem with Chaotic Encryption of voice for security, also we are using fingerprint authentication for multi-server environments to get it more secure.

Voice over Internet protocol VOIP allows us to make calls over broadband internet connection instead of regular leased lines. It converts the voice data into digital signals, divides it into packets and delivers them over respective routes and reassembles them at the receiver point. Each packet consist of source and destination addresses [1].

The most fundamental and serious thing in the transmission of voice information through IP networks, is the authentication of user identity; VoIP cannot end at physical location information to carry out certification and authentication, VoIP terminals because the physical location of information and network are independent of each other to provide for VoIP terminal with mobility. Thus, VoIP security can only be resolved through fingerprint authentication [2].

To meet the requirements of new applications with high levels of security, Chaotic Encryption with fingerprint authentication is proposed in this paper. The objective of this paper is to increase the level of security. The rest of this paper is organized as follows. Section II gives an explanation of the cryptosystem with Chaotic Encryption of speech signals. Section III gives an explanation of the fingerprint authentication. Section IV discusses the proposed technique. Section V gives the concluding remarks.

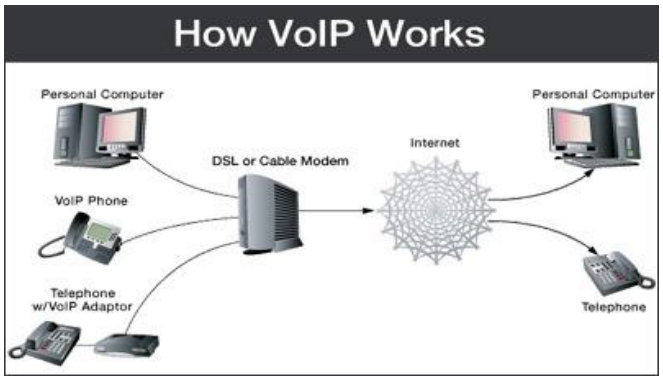


Fig. (1) VoIP.

II. CRYPTOSYSTEM WITH CHAOTIC ENCRYPTION OF SPEECH SIGNALS

This cryptosystem introduces a voice encryption approach, which is based on the permutation on speech samples with the chaotic Standard and Henon map and substitution using masks in both time and transform domains. Two parameters are taken from the main secret key to be used in the generation of the mask. DCT can be used in this cryptosystem to take out the residual intelligibility resulting from permutation and masking in the time domain encryption.

Substitution with masks is used in this cryptosystem to fill out silent periods during a voice conversation and destroy formant and pitch information. Permutation is implemented on big size blocks to grant a high degree of encryption. This cryptosystem has a low complexity, short delay, and high security. Simulation results assure that it is strong to noise [3-5].

A. Standard Map

The standard map is a two-dimension chaotic map that encrypted the file by disturbing it. It is described by the following formula:

$$\begin{bmatrix} x_{i+1} \\ y_{i+1} \end{bmatrix} = \begin{bmatrix} (x_i + y_i) \bmod N \\ \left(y_i + K \sin \frac{x_{i+1} N}{2\pi} \right) \bmod N \end{bmatrix} \quad (1)$$

B. Henon Map

A particularly simple example of a 2-dimensional map is the henon map in 1976. The henon map is one of the dynamical systems that exhibit chaotic attitudes. It is defined by two equations; the map depends on two parameters *a*, *b*, and the system exhibits a strange attractor for $a = 1.4$ and $b = 0.3$. The map repeat the point (x_n, y_n) via the equations :

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$$x_{n+1} = 1 + y_n - ax_n^2$$
$$y_{n+1} = bx_n \quad (2)$$

III. FINGERPRINT AUTHENTICATION

Is one of many biometric systems of human identification. A fingerprint sensor picks up a digital image of a fingerprint pattern, often at 500dpi (dots per inch) resolution in gray-scale using 8 bits per dot.

First, one fingerprint image (or more) is registered on the IP phone used to perform the authentication, and a template of the fingerprint is stored on a local database system.

When users want to log in, they must have their registered finger print scanned again, and a second template is generated. Next, a pattern analysis is performed to determine if there is a match and if the logon should be accepted.

The match is determined using points of interest (minutia) on the fingerprint scanned, such as ridge bifurcations and ridge endings. If enough of the minutia points and vectors match, the fingerprint scanned is considered a match and the logon is permitted [6].

A. Work flow of biometric authentication system

As shown in Fig. 2 how the biometrics authentication system works:

(1) Biometric abstraction: the raw biometric image is abstracted by the biometric scanning device.

(2) Biometric enrollment: the raw image is processed to get the measured characteristics, and the template is generated.

(3) Template storage: after the template is generated, the template is stored in memory to be verified with the input measured characteristics.

(4) Biometrics verification: the live-scanned image template is verified with the stored template.

(5) Verification result: after verification, whether the user is legal to access the system is decided [7,8].

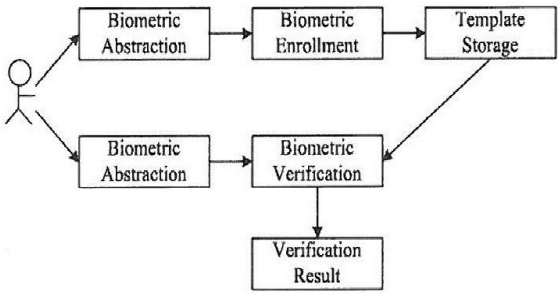


Fig. (2) How the Biometric System Works.

B. Error rates

The challenge for biometrics lies in the measurement and decision of what exactly is similar. Although biometric technology is advancing rapidly it is not yet 100% accurate in matching a previously enrolled biometric feature to a present feature. The fact is that fingerprint authentication is not 100% accurate. Inaccuracies can be caused by the condition of the finger (injured, worn, clean/dirty, wet/dry) or its presentation to the sensor (position, orientation, pressure, swiping speed). In some cases, even the user's own finger (two matching fingers) looks different to the sensor. Therefore, a biometric matching system's response is typically a matching score s (usually a single number) that quantifies the similarity between the input and the database template representations. The higher the score the more certain the system is that the two biometric measurements come from the same person [9]. A threshold (t) regulates the system decision:

- Pairs of biometric samples generating scores higher than or equal to (t) are mate pairs, they belong to the same person. The distribution of scores generated from pairs of samples from the same person is called a genuine [10].
- Pairs of biometric samples generating scores lower than (t) are non-mate pairs, they belong to different persons. The distribution of scores generated from pairs of samples from different persons is called an impostor distribution [10] as shown in Fig. 3.

The curves show false match rate (FMR) and false non-match rate (FNMR) for a given threshold t over the genuine and impostor score distributions. FMR is the percentage of non-mate pairs whose matching scores are greater than or equal to t, and FNMR is the percentage of mate pairs whose matching scores are less than t [9].

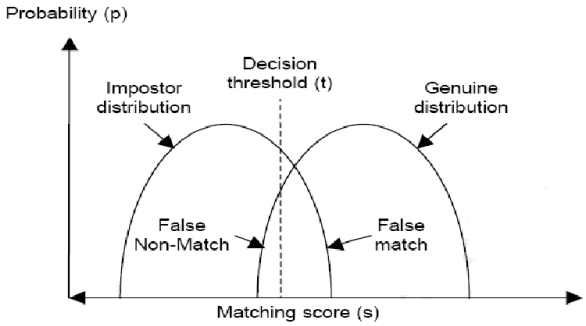


Fig. (3) Matching Score (s).

A biometric verification system can make two types of errors [11] as shown in Fig. 3 :

* False Match Rate (FMR) or False Acceptance Rate (FAR)

– The percentage of impostors the biometrics mechanism falsely authorizes. In this case two non-matching fingerprint images look similar to the authentication system.

* False Non-Match Rate (FNMR) or False Reject Rate (FRR)

– The percentage of legitimate users falsely rejected. In this case two matching fingerprint images look dissimilar to the authentication system. An operational biometric system makes a trade-off between false match rate.

(FMR) and false non-match rate (FNMR). In fact, both FMR and FNMR are functions of the system threshold: if the system's designers reduce it to make the system more tolerant to input variations and noise, FMR increases. On the other hand, if they raise it to make the system more secure, then FNMR increases accordingly [12].

Most current sensors have error rates on single

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

measurements 1% or less FMR and around 3% FNMR. That means there is less than a 1% chance that a random person can slide his finger on the sensor and be accepted [13,14].

IV. Statistical Analysis

In this paper, several metrics are selected for quality evaluation of audio cryptosystems. These metrics are Spectral Distortion, Log Likelihood Ratio, Correlation, Time analysis, Histogram and Spectogram, all are described in table I.

The SD is a form of metrics that is implemented in frequency domain on the frequency spectra of the original and processed audio signals. It is calculated in dB to show how far is the spectrum of the processed signal from that of the original signal. The SD can be calculated as follows [15,16]:

$$SD = \frac{1}{M}\sum_{m=0}^{M-1}\sum_{n=L_s m}^{L_s m+L_s-1}\left|V_s(k)-V_y(k)\right| \quad (3)$$

where $V_s(k)$ is the spectrum of the original audio signal in dB for a certain segment, $V_y(k)$ is the spectrum of the processed audio signal in dB for the same segment, M is the number of segments and L_s is the segment length. The higher the SD between the original and encrypted signals, the better is the encryption quality. On the other hand, the SD between the original signal and the decrypted signal is required to be as small as possible.

The LLR metric for an audio signal is based on the assumption that each segment can be represented by an all-pole linear predictive coding model of the form [16, 17]:

$$s(n) = \sum_{m=1}^{m_p} a_m s(n-m) + G_s u(n) \quad (4)$$

where a_m (for m=1, 2, ….., m_p) are the coefficients of the all-pole filter, G_s is the gain of the filter and $u(n)$ is an appropriate excitation source for the filter. The audio signal is windowed to form frames of 15 to 30 ms length. The LLR metric is then defined as [8]:

$$LLR = \left|\log\left(\frac{\vec{\mathbf{a}}_\mathbf{s}\overline{\mathbf{R}}_\mathbf{y}\vec{\mathbf{a}}_\mathbf{s}^\mathbf{T}}{\vec{\mathbf{a}}_\mathbf{y}\overline{\mathbf{R}}_\mathbf{y}\vec{\mathbf{a}}_\mathbf{y}^\mathbf{T}}\right)\right| \quad (5)$$

where $\vec{\mathbf{a}}_\mathbf{s}$ is the LPCs coefficient vector [1, $a_s(1)$, $a_s(2)$, . . ., $a_s(m_p)$] for the original clear audio signal, $\vec{\mathbf{a}}_\mathbf{y}$ is the LPCs coefficient vector [1, $a_y(1)$, $a_y(2)$, . . . , $a_y(m_p)$] for the decrypted audio signal, and $\overline{\mathbf{R}}_\mathbf{y}$ is the autocorrelation matrix of the decrypted audio signal. The closer the LLR to zero, the higher is the quality of the output audio signal.

Correlation coefficient between similar samples in the clear and the encrypted signals is a useful metric to assess the encryption quality of an audio cryptosystem. It can be calculated as follows:

$$r_{xy} = \frac{c_v(x,y)}{\sqrt{D(x)}\sqrt{D(y)}} \quad (6)$$

where $c_v(x,y)$ is the covariance between the original signal *s* and the encrypted signal *y*. $D(x)$ and $D(y)$ are the variances of the signals *x* and *y*, respectively. In numerical computations, the following discrete formulas can be used [18]:

$$E(x) = \frac{1}{N_x}\sum_{n=1}^{N_x} x(n) \quad (7)$$

$$D(x) = \frac{1}{Nx}\sum_{n=1}^{Nx}(x(n)-E(n))^2 \quad (8)$$

$$c_v(x,y) = \frac{1}{N_x}\sum_{n=1}^{Nx}(x(n)-E(x))(y(n)-E(y)) \quad (9)$$

where N_x is the number of audio samples involved in the calculations. The low value of the correlation coefficient rxy indicates a good encryption quality.

The processing time is the time required to encrypt/decrypt data. The smaller the processing time, the higher the speed of encryption. We have tested the proposed technique and estimated the decryption time as both the encryption and decryption processes have approximately the same time.

TABLE I
CHAOTIC MEASUREMENTS.

CHAOTIC Statistical Analysis	Encrypted	Decrypted
Spectral Distortion	21.9008	6.4823e-003
Likelihood Ratio	0.6220	4.8959e-009
Correlation	0.0051	1.0000
Processing time	0.2902	

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

V. The Proposed Technique

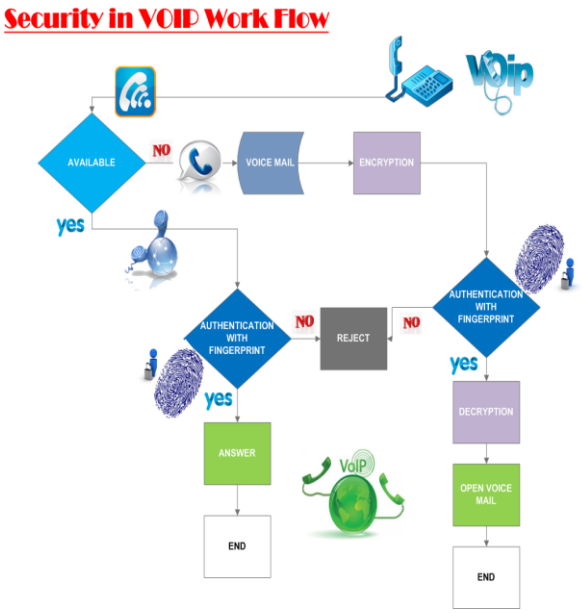


Fig. (4) Security in VoIP work flow.

The proposed technique is based on adding a stage which is fingerprint authentication to the chaotic encryption for the voice to be more secure. This stage can be performed numerically to avoid the complexity of the implementation. The second stage is the fingerprint authentication. Figure (4) show the encryption and decryption processes of the proposed technique, respectively.

With this proposed implementation, we can achieve the following gains:

- Cracking or hacking the encrypted voice becomes harder. Let us imagine the case when a hacker may crack the chaotic encryption key, i.e. the second stage, he still can not obtain the target voice as it is protected by the fingerprint authontication key.
- All acts of piracy on the encrypted voice could affect the chaotic randomized pixels. In this case, we can easily notice if the received voice has been intercepted or modified.

VI. Conclusions

A new voice cryptosystem is proposed and analyzed using several tests. Security analysis experimental results show that this cryptosystem will be very practical. This system has multi levels of security because encryption is performed with permutation process depend completely on the secret key so small change in secret key length or value will give great differ in mechanism. This system has a very powerful diffusion mechanism (a small change in clear signal makes a large change in cipher signal). The cryptosystem is tested under noise attack and it is found that it is suitable for a noisy environment.

Biometrics-based authentication has lots advantages comparing to traditional password authentication systems. Low fingerprint sensor prices, easy availability of cheap computing power and relatively good understanding of individuality information in fingerprints (compared to other biometrics) raised lot of commercial interest in fingerprint-based personal identification. As a result, many fingerprint identification vendors appeared in the last few years. Embedded applications of fingerprint-based identification (e.g., in Laptops) are on the market already.

The implementation of the proposed technique is simple, and achieves good permutation and diffusion mechanisms in a reasonable time with large immunity to noise, which is a required property for communication applications.

References

[1] Chris Roberts, “Voice Over IP”, March 2005.

[2] Liancheng Shan, Ning Jiang “Research on security Mechanism of SIP-based VoIP system” Ninth International Conference on Hybrid Intelligent Systems, 2009.

[3] Jakimoski, G. and L. Kocarev, "Chaos and Cryptography: Block Encryption Ciphers Based on Chaotic Maps", IEEE Transactions On Circuits And Systems-I: Fundamental Theory And Applications, vol. 48, no. 2, pp. 163-169, 2001.

[4] Xin Zhang, Weibin Chen, "New Chaotic Algorithm for Image Encryption", IEEE ICALIP2008, pp. 889-892, 2008.

[5] Xiping He Qionghua Zhang, "Image Encryption Based on Chaotic Modulation of Wavelet Coefficients", Congress on IEEE Image and Signal Processing (CISP'08), Sanya, Hainan, Vol. 1, pp. 622-626, May 2008.

[6] Daniel Harris, “Fingerprint Authentication”, Jun 2007.

[7] A. K. Jain and S. Pankanti: "Automated Fingerprint Identification and Imaging Systems" Advances in Fingerprint Technology, 2nd Edition, H. C. Lee and R. E. Gaensslen (eds.), Elsevier Science, 2001.Elsevier Science, 2001. URL: www.research.ibm.com/ecvg/pubs/sharat-forensic.pdf

[8] American National Standard for Information Systems. Data format for the interchange of fingerprint information URL: http://www.itl.nist.gov/iad/894.03/fing/slides/IAFIS_Overview/

[9] http://biometrics.cse.msu.edu/j2033.pdf

[10] http://www.upek.com/promlit/pdf/fltcs3a-0903.pdf

[11] SANS Security Essentials Version 2.2. Defense-In-Depth. Page 160.

[12] http://www.biometrics.org/html/bc2002_sept_program/Grother_9_02.pdf

[13] UPEK: “PerfectMatch -- Fingerprint template extraction and matching” URL:http://www.upek.com/techno/techpm.htm

[14] UPEK: “The two processes of a typical biometric application” URL:http://www.upek.com/techno/biom.htm

[15] P. Hedelin, F. Norden, F and J. Skoglund, "SD optimization of spectral coders", IEEE Workshop on Speech Coding Proceedings, pp. 28 – 30, 1999.

[16] W. Yang, M. Benbouchta, R. Yantorno, "Performance of the Modified bark Spectral Distortion as an Objective Speech Quality Measure", Proc. IEEE International Conference on Acoustic, Speech and Signal Processing, vol. 1, Washington, USA , pp. 541-544. 1998.

[17] P. Hedelin, F. Norden, F and J. Skoglund, "SD optimization of spectral coders", IEEE Workshop on Speech Coding Proceedings, pp. 28 – 30, 1999.

[18] Y. Wu and B. P. Ng, "Speech scrambling with Hadamard transform in frequency domain", Proc. 6th Int. Conf. on Signal Processing, vol. 2, pp. 1560-1563, 2002.

Biographies

El sayed Alshamy was graduated from the Faculty of electronic Engineering, Menouf University,Egypt in 1998. He is MSc. Student. He is interested in security over wired and wireless networks and speech encryption.

Performance Comparison of Column Hybrid Row Hybrid and full Hybrid Wavelet Transform on Image compression using Kekre Transform as Base Transform

Dr. H.B. Kekre
Sr. Professor, MPSTME,
Department of Computer Engineering
NMIMS University
Mumbai, India

Dr. Tanuja Sarode
Associate Professor
Department of Computer Engineering, TSEC
Mumbai University
Mumbai, India

Prachi Natu
Assistant Professor and
Ph. D. Research Scholar
MPSTME, NMIMS University
Mumbai, India

***Abstract*- This paper proposes color image compression using hybrid wavelet transform. Hybrid wavelet transform is generated using two different orthogonal component transforms. Kekre transform is used as base transform. It is combined with other transforms like DCT, DST, Hartley, Walsh, Real-DFT and Slant transform. Generated hybrid wavelet transform is applied on image as column transform, row transform and full hybrid wavelet transform to compress the image. Root mean square error is computed in these three cases at various compression ratios and compared. It has been observed that Full hybrid wavelet transform gives error which is nearly half of the error generated in column and row hybrid wavelet transform. DKT-DCT hybrid wavelet shows least RMSE than DKT-DST, DKT-Hartley, DKT-Walsh, DKT-Real-DFT and DKT-Slant hybrid wavelet transform. DKT-DCT full hybrid wavelet transform gives RMSE 11.17 at compression ratio 32, whereas using column hybrid wavelet error value is 27.85 at the same compression ratio. It shows that to save number of computations, column hybrid wavelet transforms or row wavelet transform can be used at the cost of compression ratio.**

Keywords- Hybrid Wavelet Transform; Real DFT; Kekre Transform; Image Compresion; Bit Rate

I. INTRODUCTION

In recent years image compression has become popular area for research. Large amount of data which is being transferred over internet contains multimedia data. Compression of images reduces time required to transfer this data and also saves memory space required to store it. Image compression comes under the category of lossy image compression. In lossy image compression some level of degradation of image quality takes place. But it is not perceptible to human eye and hence acceptable [1].Discrete Cosine Transform, popularly known as DCT [2] is widely used in image compression. JPEG coders based on DCT have achieved high popularity in image compression application. Wavelet transform is another technique that has gained immense popularity in recent years in many image processing applications. Image compression [3], biometrics [4,5,6], image segmentation, content based image retrieval [7], steganography [8] are to name the few. Basically wavelets are mathematical tools used to extract information from different kind of data.

Concept of wavelet was first introduced by Jean Morlet in 1982[9]. Wavelets have high energy compaction property. It allows achieving higher compression ratio in image compression applications. Wavelet transform is good alternative to short time Fourier transform (STFT) [10,11]. Unlike STFT, wavelet transforms use variable window sizes that change along the frequency range [12]. Traditional study of wavelet shows that Daubechies wavelet [13] and Haar wavelets have been used and analyzed for image compression application. Latest research work focuses on Walsh wavelet, Hartley Wavelet, Kekre Wavelet and slant Wavelet transforms which are generated by Kekre's algorithm[14]. Motive of this paper is to propose image compression using hybrid wavelet transform and compare the performance of column hybrid wavelet, row hybrid wavelet and full hybrid wavelet in image compression application. Remaining sections of this paper are organized as

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

follows: Section II contains related work in the area of image compression. Proposed method is discussed in section III. Results based on experimental work are discussed in section IV. Finally, section V concludes the work done.

II. Related Work

Large amount of research has been done on image compression and still is in progress. Performance of different wavelets in image compression has been compared by M.S. Abdullah and N. Subba Rao [15]. It compares performance of Haar wavelet, Daubechie wavelet, Coiflet wavelet, Biorthognal wavelet, Demeyer wavelet, and Symlet wavelet. Also lifting based wavelet transforms and Set Partitioning in Hierarchical Trees (SPIHT) algorithm are used to compress the input images. Taubman [16] developed adaptive wavelet transforms to modify the prediction step by using the properties of the image. Calypoole et al. [17] have proposed a prediction operator based on the local properties of the image. In [18] Omar et al. described an adaptive polyphase structure based on the reduction of the variance. In [19] Boulgouries et al. have calculated the optimal predictors, by minimizing the prediction error variance, and then these optimal predictor filters have been applied with adaptive update filters. Image compression using Daubechies 9/7 wavelet has been proposed by Qiu-yuan et al.[20] where three level Wavelet transform has been applied on grayscale image. PSNR up to 38.76 is obtained using this method. Use of Kekre wavelet, Slant wavelet and Walsh wavelet is proposed in [21] by H.B. Kekre, Tanuja Sarode and Prachi Natu. These wavelet transforms are generated using the algorithm in [22] which has been proposed by Kekre et.al. Flexibility to vary the component size is provided [14] to choose the best component size. Real Fourier transform which contains real coefficients of Fourier transform is introduced in [23] by Kekre et al. It does not contain complex exponentials like in Fourier transform. Quantization technique followed by run length coding is combined to achieve image compression by Samir Kumar et.al. in[24]. It is effective for images with large similar locality of pixel lay out. New trend of using hybrid techniques for image compression is in now. Combination of wavelet transform and neural network has been proposed by M. Venkata Subbarao et al. [25]. Use of neural network with wavelets helps to get higher compression ratio. Various other hybrid techniques are available in literature. This paper focuses on Hybrid Wavelet Transform for image compression and compares performance of column hybrid wavelet and Row hybrid wavelet with full Hybrid Wavelet transform.

III. Proposed Method

In this paper, image compression is achieved using hybrid wavelet transform. Hybrid wavelet transform is generated using two different orthogonal transforms. To generate hybrid wavelet of A and B, first 'm' rows of resultant matrix is calculated by repeating each column of 'A' 'N' times and multiplying it with each element of first row of B. These 'M' rows represent global characteristics in hybrid wavelet transform. Remaining rows are obtained by translating the rows of matrix B from second row onwards. These rows contribute local features of an image. Generated hybrid wavelet transform is shown below in Fig.1. As shown in Fig. 1 Kekre Transform (DKT) plays role of matrix 'A'. As selection of Kekre transform as base transform gives acceptable image quality at higher compression ratios, than any other transform as a base transform, it is selected as matrix 'A'. Hybrid wavelet HAB is generated by combining DKT with other orthogonal transforms like DCT, DST, Walsh, Slant, Hartley and Real-DFT transform. Using these hybrid wavelet transforms, transformed image is obtained as follows:

f= original image
H_{AB}= Hybrid Wavelet Transform
F= Transformed Image

$$[F]= [H_{AB}]*[f] \text{ gives column transformed image} \quad (1)$$

$$[F]= [f]*[H_{AB}]^T \text{ is row transformed image} \quad (2)$$

$$[F]= [H_{AB}]*[f]*[H_{AB}]^T \text{ is full transformed image} \quad (3)$$

Inverse Column transform is obtained as:

$$[f]= [F]*[H_{AB}]^T \quad (4)$$

Inverse Row Transform is

$$[f]= [H_{AB}]*[F] \quad (5)$$

And Inverse Transform for Full Hybrid Wavelet is

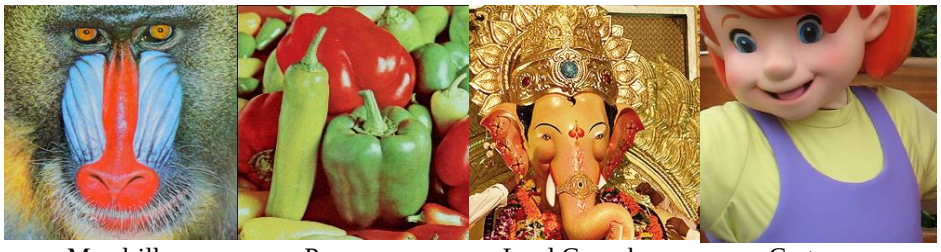
$$[f]= [H_{AB}]^T*[F]*[H_{AB}] \quad (6)$$

Column transform of individual red, green and blue plane of an image is calculated using eq. 1. Similarly row and full transform is also applied on individual plane using respective equations mentioned above. High frequency coefficients are eliminated from transformed image to obtained compressed image. Compression ratio is varied up to 32 and root mean square error is calculated for various compression ratios. This procedure is followed for row transform and full transform of images.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$b_{11}\begin{pmatrix}a_{11}\\a_{21}\\.\end{pmatrix}$	…	$b_{1n}\begin{pmatrix}a_{11}\\a_{21}\\.\end{pmatrix}$	$b_{11}\begin{pmatrix}a_{12}\\a_{22}\\.\end{pmatrix}$	…	$b_{1n}\begin{pmatrix}a_{12}\\a_{22}\\.\end{pmatrix}$	….	$b_{11}\begin{pmatrix}a_{1m}\\a_{2m}\\.\end{pmatrix}$	…..	$b_{1n}\begin{pmatrix}a_{1m}\\a_{2m}\\.\end{pmatrix}$
b_{21}	…	b_{2n}	0	0	0	…..	0	…	0
0	0	0	b_{21}	…	b_{2n}	…..	0	…	0
. .	. .	. .	. .	. .	. .	…..	. .	. .	. .
0	0	0	0	0	0	.	b_{21}	…	b_{2n}
b_{31}	…	b_{3n}	0	0	0	.	0	…	0
0	0	0	b_{31}	…	b_{3n}	.	0	…	0
. .	. .	. .	. .	. .	. .	.	. .	. .	. .
0	0	0	0	0	0	.	b_{31}	…	b_{3n}
. .	. .	. .	. .	. .	. .	.	. .	. .	. .
b_{n1}	…	b_{nn}	0	0	0	.	0	0	0
0	0	0	b_{n1}	…	b_{nn}	.	0	0	0
. .	. .	. .	. .	. .	. .	.	. .	. .	. .
0	0	0	0	0	0	.	b_{n1}	…	b_{nn}

Fig.1. Generation of Hybrid Wavelet Transform from Two Orthogonal Component Transform



Mandrill Peppers Lord Ganesha Cartoon

Dolphin Waterlili Bud Bear

Lena Lemon Ball Balloon

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

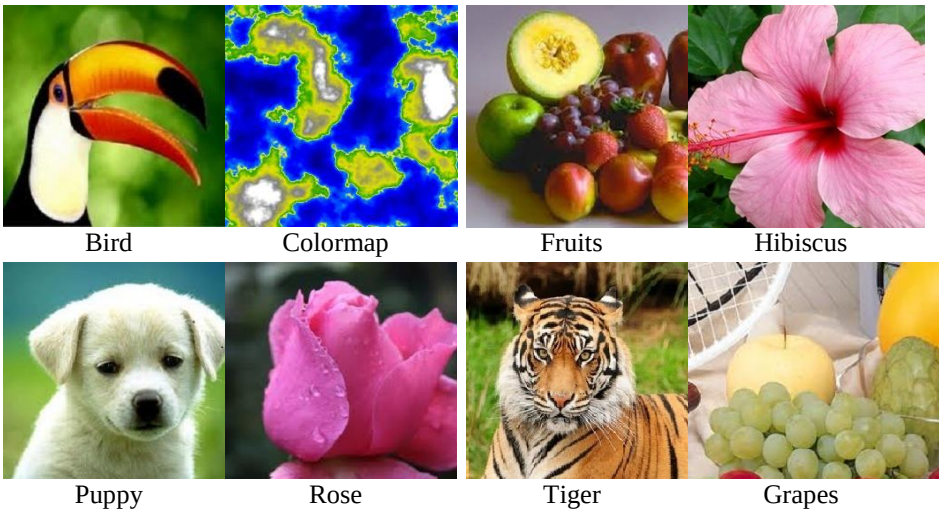


Fig. 2 Set of color images use for experimental purpose

IV. Experimental Work and Results

Proposed method is applied on set of twenty color images. Images selected for experimental work are shown below in Fig. 2. Execution is performed using Matlab 7.0 on AMD Dual core processor with 4 GB RAM.

Column transform is applied to the images using eq. (1). DKT-DCT pair is selected for testing. Size of component transform DKT (base transform) is denoted as 'm' and local transform size is denoted as 'n'. Using different values of m and n, hybrid wavelet of size 256x256 is generated and applied to the images. In each case, root mean square error between original image and compressed image is calculated. Graph of error against compression ratio is plotted in Fig. 3.

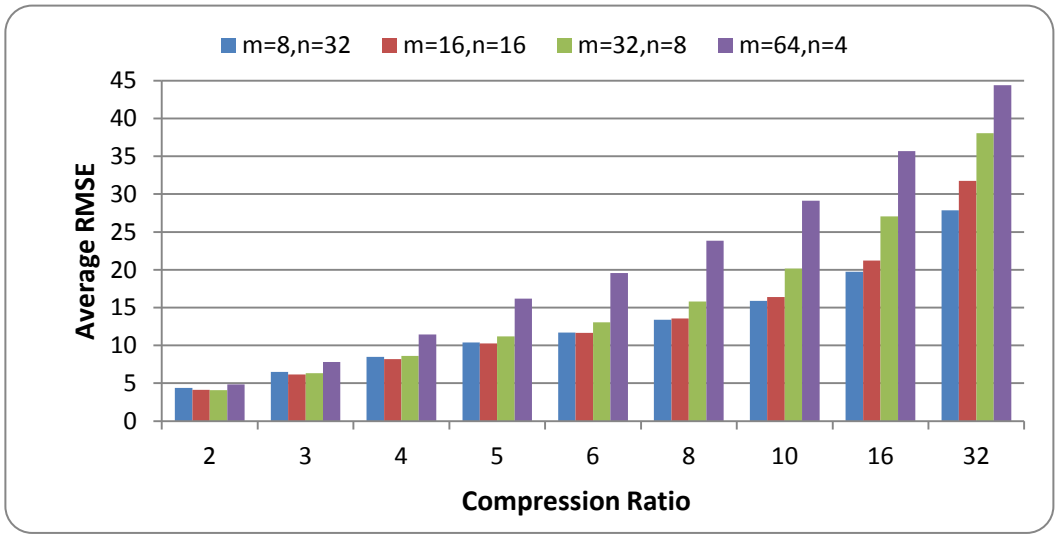


Fig. 3 Average RMSE against Compression Ratio using different component sizes of DKT-DCT Column Hybrid Wavelet Transform

It shows that, among 4 different size combinations of m and n, m=8 and n=32 (i.e. 8-32) size of component transforms gives lower RMSE value. At compression ratio 16, using component size m=8 and n=32, RMSE obtained is 19.73. It gives acceptable quality of reconstructed image. At compression ratio 32, RMSE rises to 27.85 affecting the quality of reconstructed image.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Fig. 4 compares RMSE obtained using DKT-DCT Row Hybrid Wavelet Transform. Different component size like 8-32, 16-16, 32-8, 64-4 is selected and RMSE obtained in each case is observed for compression ratios from 2 to 32. At compression ratio 16, RMSE calculated using 8-32 component size is 20.01which is slightly higher than RMSE obtained using DKT_DCT column hybrid wavelet transform. It gives visually acceptable image. At compression ratio 32 RMSE increases up to 27.89. But at this higher compression ratio reconstructed image quality deteriorates. In both column and row hybrid wavelet transform component size 8x8 and 32x32 gives lower RMSE. Hence this size is selected to observe the performance of other column and row hybrid wavelet transforms formed with DKT as a base transform.

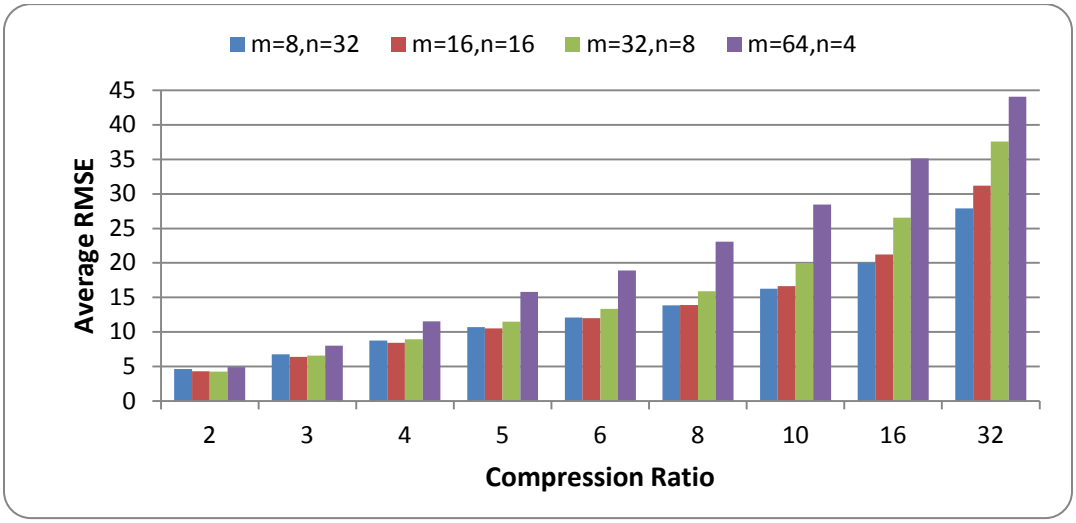


Fig. 4 Average RMSE against Compression Ratio using different component sizes of DKT-DCT Row Hybrid Wavelet Transform

Fig. 5 compares the performance of various column hybrid wavelet transforms generated using component transforms DKT (8x8) and local component transform (32x32). DCT, Slant, Hartley, Real-DFT, DST and Walsh transform are used to represent local features of image. DKT-DCT column hybrid wavelet transform gives lesser value of RMSE than others which is followed by DKT-Slant pair. Using column hybrid wavelet satisfactory image quality is obtained even at compression ratio 16.

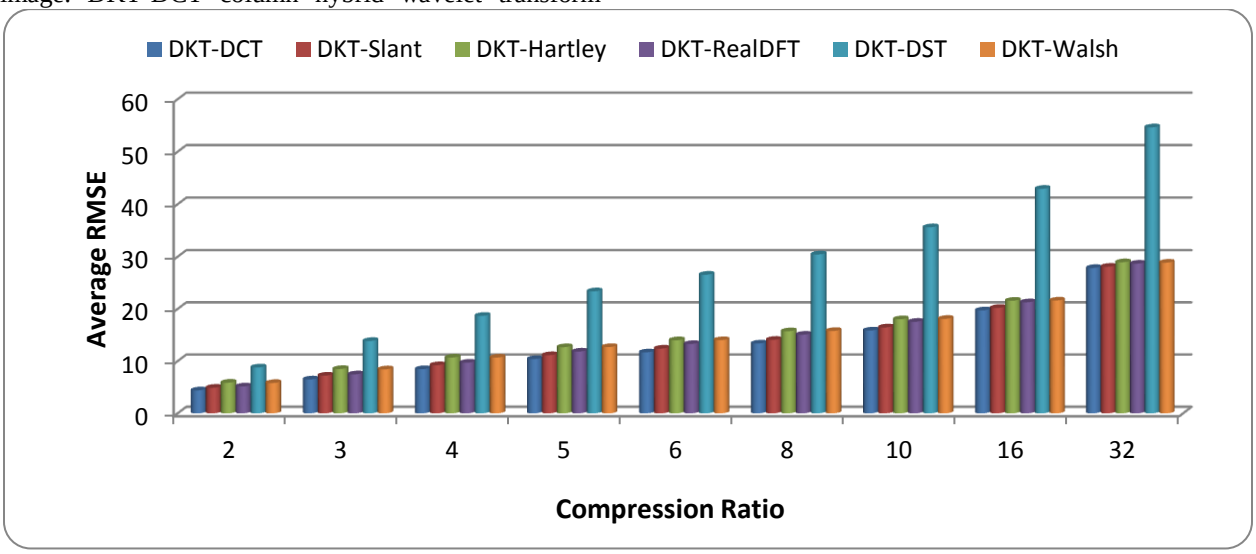


Fig.5. Avg. RMSE against Compression Ratio using DKT of size 8x8 and local component size 32x32 in Column hybrid wavelet Transform

Fig.6 shows comparison of RMSE in row hybrid wavelet transforms against compression ratio. Like column hybrid wavelet, here also DKT-DCT performs better than all other row hybrid wavelets. Performances

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

of all row hybrid wavelets are nearly equal except DKT-DST. Using Row hybrid wavelet transform, 16:1 compression can be achieved.

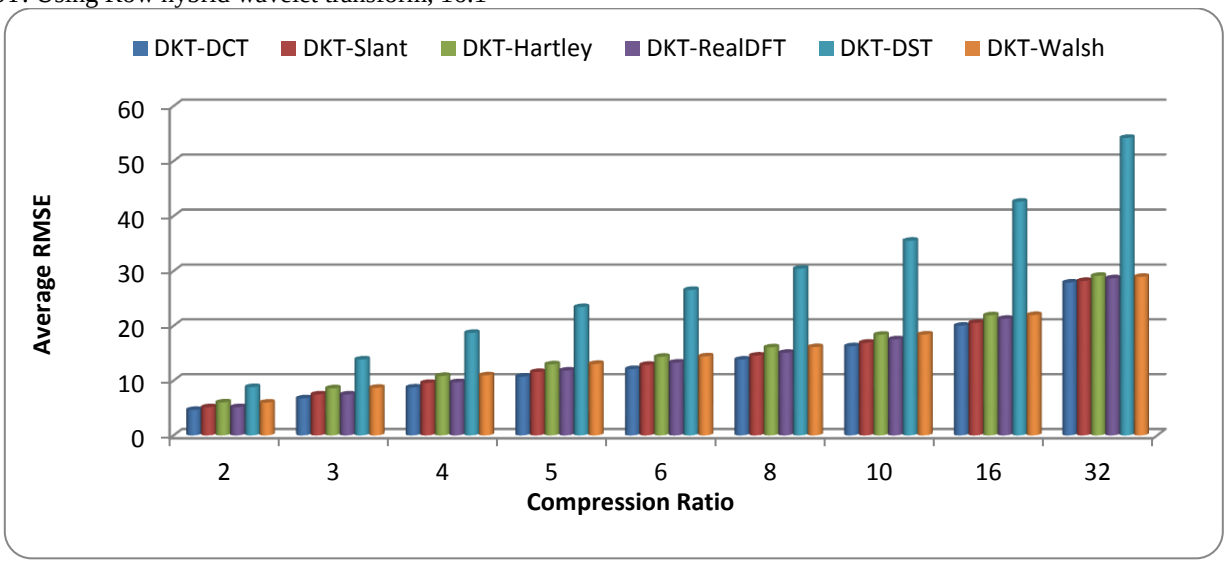


Fig. 6 Avg. RMSE against Compression Ratio using DKT of size 8x8 and local component size 32x32 in Row hybrid wavelet Transform

Graph in Fig. 7 compares RMSE computed in DKT-DCT column hybrid, row hybrid and full hybrid wavelet transform. It has been observed that performance of column hybrid wavelet and row hybrid wavelet is nearly equal. Full hybrid wavelet transform reduces RMSE to half than RMSE in column and Row Hybrid Wavelet transform. At compression ratio 32, RMSE given by DKT-DCT column hybrid is 27.85, in Row hybrid Wavelet it is 27.89 which is almost same as in column hybrid wavelet. In DKT-DCT full hybrid wavelet error drastically reduces to 11.17. It reduces by more than 50% as that of in column hybrid and row hybrid wavelet transform.

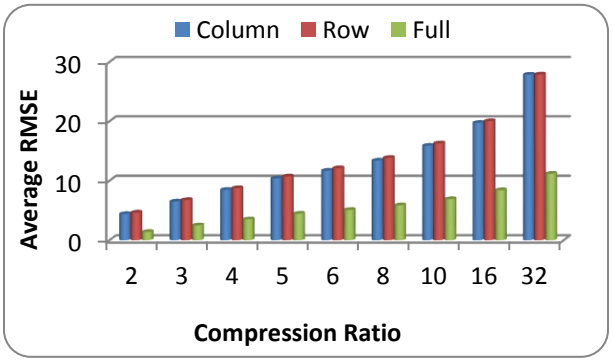


Fig.7. Average RMSE against Compression Ratio using component size of DKT=8x8 and DCT=32x32 Row, Column and Full Hybrid Wavelet Transform

Fig. 8 plots RMSE against Bit Rate for DKT-DCT Hybrid Wavelet Transform. It compares RMSE in column, row and full hybrid wavelet. As less number of bits is used to represent the image, more compression is obtained. Minimum bit rate is of 0.25 bpp is possible in full hybrid wavelet. At this bit rate fine image quality is obtained with error 11.17. Row and column hybrid wavelet show equal performance.

Fig. 9 shows reconstructed images in DKT-DCT Hybrid Wavelet transform. All three cases i.e. column wavelet, Row wavelet and full Wavelet are considered. From these images we can say that in full hybrid wavelet good quality of reconstructed image is obtained. RMSE in full hybrid wavelet is one third as compared to reconstructed image in column hybrid wavelet transform and hence image quality is much better even at high compression ratio 32. Using column hybrid wavelet transform acceptable image quality is obtained at compression ratio 16. Beyond this compression ratio RMSE increases. It results in degradation in quality of compressed image. But to save the computational overhead and to achieve reasonable image quality column hybrid wavelet can be used instead of full hybrid wavelet up to compression ratio 16.

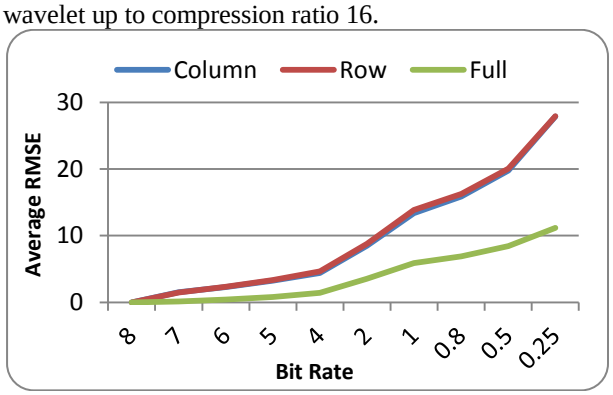


Fig.8. Average RMSE against Bit Rate using component size of DKT=8x8 and DCT=32x32 Row, Column and Full Hybrid Wavelet Transform

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Reconstructed Images at different Compression Ratios using DKT-DCT Column Hybrid Wavelet Transform						
Compression Ratio						
Original Image	2	4	8	10	16	32
						
RMSE→	1.88	3.77	6.06	7.46	9.82	17.69
Reconstructed Images at different Compression Ratios using DKT-DCT Row Hybrid Wavelet Transform						
Compression Ratio						
Original Image	2	4	8	10	16	32
						
RMSE→	2.07	3.86	6.76	8.08	10.62	17.19
Reconstructed Images at different Compression Ratios using DKT-DCT Full Hybrid Wavelet Transform						
Compression Ratio						
Original Image	2	4	8	10	16	32
						
RMSE→	0.32	0.95	1.94	2.46	3.30	4.85

Fig.9. Reconstructed images using DKT-DCT column Hybrid Wavelet, Row Hybrid Wavelet and full Hybrid Wavelet Transform at various compression ratios with respective RMSE values

Fig. 10 shows reconstructed images using column, row and full hybrid wavelet transform of DKT-Walsh. In column and row hybrid wavelet compression ratio of 16 can be achieved with moderate image quality. Image is distorted at compression ratio 32 with RMSE 18.75 and 18.27 respectively. In full hybrid wavelet reconstructed image is superior with RMSE 12.58. Sample decompressed image in DKT-Slant Hybrid wavelet is shown in Fig.11. Using full hybrid wavelet transform clear decompressed images are obtained at compression ratio 32 showing error value 5.57. Using column and row hybrid wavelet error increases three times than full wavelet. Hence acceptable compression ratio is 16 in this case. Decompressed images in DKT-Hartley hybrid wavelet are shown in Fig. 12 and Fig. 13 shows sample images using DKT-RealDFT hybrid wavelet transform. Their results are similar to DKT-Slant i.e. up to compression ratio 16, results of column and row hybrid wavelet are acceptable. To achieve higher compression ratio full hybrid wavelet transform must be selected.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

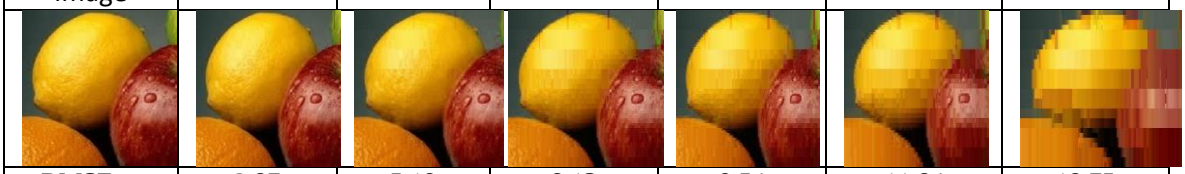
Reconstructed Images at different Compression Ratios using DKT-Walsh Column Hybrid Wavelet Transform						
	Compression Ratio					
Original Image	2	4	8	10	16	32
						
RMSE→	2.87	5.19	8.13	9.54	11.84	18.75
Reconstructed Images at different Compression Ratios using DKT-Walsh Row Hybrid Wavelet Transform						
	Compression Ratio					
Original Image	2	4	8	10	16	32
						
RMSE→	2.87	5.58	8.46	10	12.62	18.27
Reconstructed Images at different Compression Ratios using DKT-Walsh Full Hybrid Wavelet Transform						
	Compression Ratio					
Original Image	2	4	8	10	16	32
						
RMSE→	0.54	1.99	4.53	5.92	8.14	12.58

Fig.10. Reconstructed images using DKT-Walsh column Hybrid Wavelet, Row Hybrid Wavelet and full Hybrid Wavelet Transform at various compression ratios with respective RMSE values

Reconstructed Images at different Compression Ratios using DKT-Slant Column Hybrid Wavelet Transform						
	Compression Ratio					
Original Image	2	4	8	10	16	32
						
RMSE→	2.39	4.22	6.91	7.92	10.19	17.88
Reconstructed Images at different Compression Ratios using DKT-Slant Row Hybrid Wavelet Transform						
	Compression Ratio					
Original Image	2	4	8	10	16	32

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

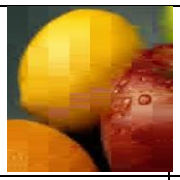
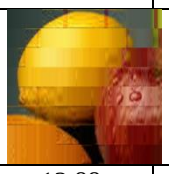
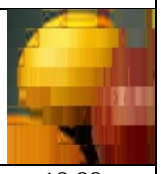
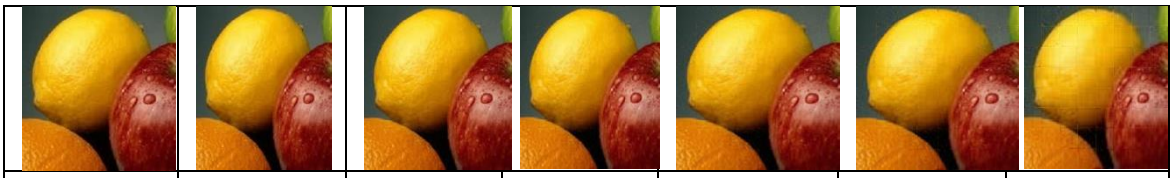
						
RMSE→	2.38	4.40	7.17	8.47	10.97	17.43
Reconstructed Images at different Compression Ratios using DKT-Slant Full Hybrid Wavelet Transform						
Compression Ratio						
Original Image	2	4	8	10	16	32
						
RMSE→	0.48	1.40	2.64	3.21	4.05	5.57

Fig.11. Reconstructed images using DKT-Slant column Hybrid Wavelet, Row Hybrid Wavelet and full Hybrid Wavelet Transform at various compression ratios with respective RMSE values

Reconstructed Images at different Compression Ratios using DKT-Hartley Column Hybrid Wavelet Transform						
Compression Ratio						
Original Image	2	4	8	10	16	32
						
RMSE→	3.31	6.13	9.18	10.60	12.88	18.88
Reconstructed Images at different Compression Ratios using DKT-Hartley Row Hybrid Wavelet Transform						
Compression Ratio						
Original Image	2	4	8	10	16	32
						
RMSE→	3.27	6.23	9.27	10.67	13.14	18.89
Reconstructed Images at different Compression Ratios using DKT-Hartley Full Hybrid Wavelet Transform						
Compression Ratio						
Original Image	2	4	8	10	16	32

http://sites.google.com/site/ijcsis/
ISSN 1947-5500



RMSE→	0.46	1.39	2.84	3.58	4.77	7.07

Fig.12. Reconstructed images using DKT-Hartley column Hybrid Wavelet, Row Hybrid Wavelet and full Hybrid Wavelet Transform at various compression ratios with respective RMSE values

Reconstructed Images at different Compression Ratios using DKT-RealDFT Column Hybrid Wavelet Transform

Original Image	Compression Ratio 2	4	8	10	16	32
						
RMSE→	2.51	4.78	7.85	9.36	11.71	18.72

Reconstructed Images at different Compression Ratios using DKT-Real DFT Row Hybrid Wavelet Transform

Original Image	Compression Ratio 2	4	8	10	16	32
						
RMSE→	2.40	5.06	8.17	9.75	12.46	18.34

Reconstructed Images at different Compression Ratios using DKT-Real DFT Full Hybrid Wavelet Transform

Original Image	Compression Ratio 2	4	8	10	16	32
						
RMSE→	0.34	1.06	2.23	2.86	3.87	5.86

Fig.13. Reconstructed images using DKT-RealDFT column Hybrid Wavelet, Row Hybrid Wavelet and full Hybrid Wavelet Transform at various compression ratios with respective RMSE values

Figure 14 shows images obtained using DKT-DST hybrid wavelet. Image quality degrades from compression ratio 8 onwards in column and row hybrid wavelet transform. Blocking effect is observed in full hybrid wavelet at compression ratio 16. It becomes more prominent as compression ratio increases to 32.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

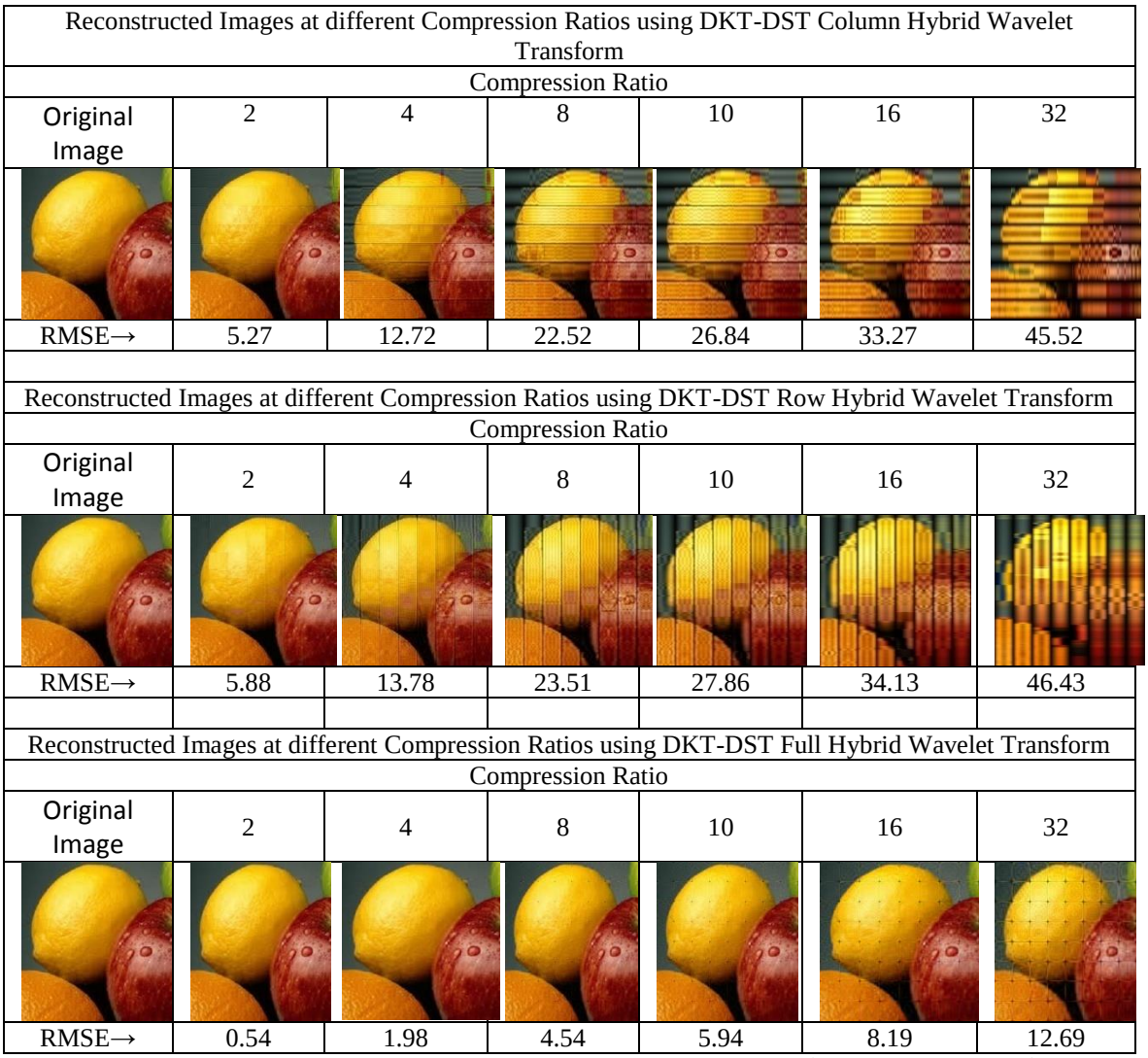


Fig.14. Reconstructed images using DKT-DST column Hybrid Wavelet, Row Hybrid Wavelet and full Hybrid Wavelet Transform at various compression ratios with respective RMSE values

V. CONCLUSION

This paper compares the performance of column, row and full hybrid wavelet transform on color image compression. Using MxM and NxN size orthogonal component transforms, hybrid wavelet transform of MNxMN size is generated. Different values of M and N can be selected to generate MNxMN hybrid wavelet. Here 256x256 hybrid wavelet is generated by varying value of M as 8, 16, 32 and 64. Kekre transform is chosen as base transform and local component transforms are varied as DCT, DST, Walsh, Hartley, Real-DFT and Slant. Among these different combinations DKT-DCT full hybrid wavelet transform gives lower RMSE 11.17 at compression ratio 32 as compared to other hybrid combinations. Component transform size that gives this minimum value is M=8 and N=32. DKT-Slant, DKT-Hartley, DKT-Walsh and DKT-Real-DFT show nearly equal performance.

Column or row hybrid wavelet can also be used for compression. It reduces computational overhead to half than in full hybrid wavelet transform but at the cost of compression ratio. Using column transform maximum compression ratio up to 16 is obtained. In DKT-DCT column hybrid wavelet transform, RMSE obtained at compression ratio 16 is 19.73. Further increase in compression ratio degrades the quality of reconstructed image. Column and row hybrid wavelet transforms show nearly same performance. Using DKT-DCT least number of bits per pixels can be used to represent the image. Possible bpp 0.25 gives error value 11.17 for DKT-DCT full hybrid wavelet resulting in fine image quality.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

REFERENCES

[1] Ilker Hacihaliloglu, Mesat Kartal, "DCT and wavelet Based Image compression in Satellite Images", in proc. of IEEE conference on Recent Advances on Space Technologies, 2003, pp. 79-84.

[2] Ahmed, N., Natarajan T., Rao K. R.: Discrete cosine transform. In: IEEE Transactions on Computers, Vol. 23, 90-93, 1974.

[3] H.B.Kekre, Tanuja Sarode, Prachi Natu, "Image Compression using Column, Row and Full Wavelet Transforms of Walsh, Cosine, Haar, Kekre, Slant and Sine and their Comparison with Corresponding Orthogonal Transforms", International Journal of Engineering Research and Development(IJERD)**,** Volume 6, Issue 4 (March 2013), PP.102-113.

[4] H.B. Kekre, Dr, Tanuja Sarode, Prachi Natu, "Performance Comparison of face Recognition using DCT and Walsh Transform with Full and Partial Feature Vector Against KFCG VQ Algorithm", In proc.of 2nd International Conference and workshop on Emerging Trends in Technology (ICWET) 2011 published in International Journal of Computer Applications (IJCA), 2011, pp.22-30.

[5] H. B. Kekre, Dr, Tanuja Sarode, Prachi Natu, "Speaker identification using 2D DCT, Walsh and Haar on full and block Spectrograms", International Journal of Computer Science and Engineering, (IJCSE), Volume 2, Issue 5, 2010.

[6] H.B. Kekre, Tanuja Sarode, Rekha Vig, " Multi-resolution Analysis of Multi-spectral Palmprints using Hybrid Wavelets for Identification", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 4, No.3, 2013.

[7] H. B. Kekre, Dhirendra Mishra, " Image Retrieval using DST and DST Wavelet Sectorization", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 2, No. 6, 2011, pp. 91-97.

[8] H.B.Kekre, Archana Athawle, "Information Hiding using LSB Technique with Increased Capacity", International Journal of Cryptography and Security, Vol.1, No. 2, Oct 2008.

[9] M. Sifuzzaman, M.R. Islam and M.Z. Ali "Application of Wavelet Transform and its Advantages Compared To Fourier Transform", Journal of Physical Sciences, Vol. 13, 2009, pp. 121-134.

[10] Sonja Grgic, Kresimir Kers, and Mislav Grgic, "Image compression using Wavelets," ISIE 1999-Bled, Slovenia, pp. 99-104.

[11] Julius O. Smith III and Xavier Serra," An Analysis/Synthesis Program for Non-Harmonic Sounds Based on a Sinusoidal Representation", Proceedings of the International Computer Music Conference (ICMC-87, Tokyo), Computer Music Association, 1987.

[12] Olivier Rioul and Martin Vetterli, "Wavelets and Signal Processing," IEEE Signal Processing Magazine, pp.14-28, October 1991.

[13] Daubechies, I. "The wavelet transform, time-frequency localization and Signal analysis", IEEE Transformation and Information Theory 36: 1990, 961-1005.

[14] H.B. Kekre, Tanuja Sarode, Prachi Natu, "Performance Comparison of Walsh Wavelet, Kekre Wavelet and Slant Wavelet Transform in Image Compression", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 10, October 2013, pp. 3834-3840.

[15] M.S. Abdullah, N. Subba Rao, "Image Compression using Classical and Lifting based Wavelets", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 8, August 2013, pp. 3193-3198.

[16] D. Taubman "Adaptive non-separable lifting transforms for image compression", in Proc. Int. Conf. Image Processing, Kobe, Japan, Oct.1999.

[17] R.L. Caypoole, G.M. Davis, W. Sweldens, and R.Gboranuk, "Nonlinear wavelet transforms for image coding via lifting", IEEE Trans. on Image Processing, Vol.12, pp. 1149–1459, 2003.

[18] Omer Nezih Gerek, A.Enisçetin, "Adaptive polyphase decomposition structure for image compression", IEEE Trans. on Image Processing, Vol.9 (10), 2000.

[19] N.V. Boulgouris, Dimitrios Tzavaras and Michael Gerassiomos Strintzis, 2001, "Lossless image compression based on optimal prediction, adaptive lifting and conditional arithmetic coding", IEEE Transaction on Image Processing, Vol.10 (1), pp. 1–14.

[20] Qiu-yuan CAI, Geng-sheng WANG, Yun-xin YU, "Research of Still Image Compression Based on Daubechies 9/7 Wavelet Transform", 2nd International Conference on Future Computer and Communication pp. 357-361, 2010.

[21] H.B. Kekre, Tanuja Sarode, Prachi Natu, "Performance comparison of Walsh Wavelet, Kekre Wavelet and Slat Wavelet Transform in Image Compression", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 10, October 2013, pp. 3834-3840.

[22] H.B. Kekre, Tanuja Sarode, Sudeep Thepade, Sonal Shroff, "Instigation of Orthogonal Wavelet Transforms using Walsh, Cosine, Hartley, Kekre Transforms and Their use in Image Compression" International Journal of Computer Science and Information Security, Vol. 9, No. 6, 2011.

[23] H.B. Kekre, Tanuja Sarode, Prachi Natu, "Image Compression using Real Fourier Transform, It's Wavelet Transform and Hybrid Wavelet with DCT", Accepted in International Journal of Advanced Computer Science and Applications,(IJACSA) Vol. 4, No.5, 2013.

[24] Samir Kumar Bandyopadhyay, Tuhin Utsab Paul, Avishek Raychoudhury, "Image Compression using Approximate Matching and Run Length", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 2, No. 6, 2011, pp. 117-121.

[25] M. Venkata Subbarao, N.Sayedu Khasim, Jagadeesh Thati and M. H. H.Sastry, "Hybrid Image Compression using DWT and Neural Networks", International Journal of Advanced Science and Technology Vol. 53, April, 2013.

AUTHORS PROFILE

Dr. H. B. Kekre has received B.E (Hons.) in Telecomm Engineering from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960. He has guided 17 Ph.Ds, more than 100 M.E./M.Tech and several B.E. / B.Tech projects, while in IIT and TSEC. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 450 papers in National / International Journals and Conferences to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE, Life Member of ISTE and Senior Member of International Association of Computer Science and Information Technology (IACSIT). Recently fifteen students working under his guidance have received best paper awards. Currently eight research scholars working under his guidance have been awarded Ph. D. by NMIMS (Deemed to be University). At present eight research scholars are pursuing Ph.D. program under his guidance.

Dr. Tanuja K. Sarode has received M.E. (Computer Engineering) degree from Mumbai University in 2004, Ph.D. from Mukesh Patel School of Technology, Management and Engg. SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 11 years of Experience in teaching. Currently working as Assistant Professor in Department of

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Computer Engineering at Thadomal Shahani Engineering College, Mumbai. She is member of International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT). Her areas of interest are Image Processing, Signal Processing and Computer Graphics. She has 150 papers in National /International Conferences/journal to her credit.

Ms. Prachi Natu has received M.E. (Computer) degree from Mumbai University in 2011. Currently pursuing Ph.D. from NMIMS University. She has 08 years of experience in teaching. Currently working as Assistant Professor in Department of Computer Engineering at Mukesh Patel School of Technology, Management and Engineering. Vile Parle (Mumbai). Her areas of interest are Image Processing, Database Management Systems and Operating Systems. She has 18 papers in International Conferences/journal to her credit.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

AN OPTIMIZED MULTIPLIER ARCHITECTURE FOR A DUAL FIELD PROCESSOR FOR ECC

Nagaraja Shylashree
Research Scholar, Department of E & C Engineering
VTU, PESCE, Mandya
Karnataka, India

Venugopalachar Sridhar
Department of E & C Engineering.
VTU, PESCE, Mandya
Karnataka, India .

***Abstract*—This paper presents an optimized multiplier architecture for a Dual field Elliptic curve cryptographic (DF-ECC) processor for both Galois prime field GF(p) and binary field GF(2^m), by the selection of proper type of coordinates in the Elliptic curve cryptographic (ECC) arithmetic unit. In this unit, multiplication plays important role to perform scalar multiplication (kP) .We present two multiplier architectures, using FPGA Virtex-5 as the target device. These are, multiplier using carry save adder (CSA) and vedic multiplier. The proposed Dual field ECC processor using Vedic multiplier can reach a speed up to 151MHz.**

Keywords- Dual Field ECC; Galois Field; Elliptic curve scalar multiplication;Elliptic curve point addition;Elliptic curve point doubling; Dual field multiplier; Dual field adder;Vedic multiplier;FPGA.

I. INTRODUCTION

Security plays a vital role in Industrial applications. Normally data will be transmitted or shared through insecure channels. In order to protect the confidential data, cryptography has become very important and striving issue. Accordingly cryptography is mainly used for confidentiality, authentication, data integrity, and non-repudiation. Cryptography can be divided into two types: secret-key cryptography and public-key cryptography. Secret-key cryptography [3]-[4], which usually has a relatively compact architecture and smaller key size than public-key cryptography, is often used to encrypt/decrypt sensitive information or documents. Elliptic curve cryptography (ECC) is one of the public key crypto-system. Elliptic curve cryptography (ECC) was proposed by Koblitz [1] and Miller in 1985 [2]. Its attractive feature is smaller key size with the same level of security compared to other cryptographic algorithms like RSA. For example, the security level of 160 and 224 bits ECC is analogous to the 1024 and 2048 bits RSA respectively [15][16]. Point addition and doubling are key operations of ECC which decide the Performance of ECC. In Refs. [7],[8], architectures are proposed using parallelism and pipelining in both addition and doubling by using the projective coordinates. Scalar multiplication based on Montgomery method is proposed which reduces delay by merging addition and doubling. Multiplication in finite fields takes more time than addition and squaring. Reductions are defined within a multiplier unit to achieve high throughput. A high performance ECC processor based on the Lopez–dahab EC point multiplication was proposed [9]. A dual field ECC processor with projective coordinates adaptive to both binary and prime fields, implemented with carry save adder(CSA) as multiplier was proposed [10]. In [11] the unified multiplier architecture was proposed, which operates in both fields that is GF(p) and GF(2^m).This work[12] proposes a unified architecture for a public key cryptographic processor, that can support both RSA and ECC. In [13], the theoretical model is used to approximate the delay of critical paths in the elliptic curve scalar multiplier architecture (ECSMA) implemented on k input lookup table (LUT)-based field-programmable gate arrays. In [14] the implementations of different software algorithms to compute the scalar multiplication based on Elliptic Curve Cryptography (ECC) using five different methods are studied. These are defined for both prime and binary fields. Several improved ECC architectures have been proposed [17-26]. Among them [18],[19]&[21-24] are implemented on FPGA platform.

Compared to ASIC, FPGA has the advantage that any cryptographic processor can be re-configured or re-programmed based on the security and application requirements. Therefore, Field Programmable Gate Array (FPGA) implementation is well suited platform for cryptographic processors.

In ECC, normally two finite fields are used. Those are Binary field known as Galois field GF(2^m) and Prime field known as Galois field GF(p). In prime field GF(p) the hardware complexity will be slightly higher than the Binary field GF(2^m). Since binary field is suitable for hardware application, as on literature date, only a few hardware implementations are done in prime field GF(p).

The organization of this paper is as follows: Mathematical background is defined in Section II. Sections III and IV discuss about CSA and Vedic mathematics. The proposed multiplier architecture is structured in Section V. Section VI presents the Dual field ECC processor. The Experimental results and comparisons are given in Section VII. Finally Section VIII concludes the work.

II. BACKGROUND FOR ECC

As described in the Specifications for public key cryptography mentioned by IEEE 1363 standard [5] and recommended by NIST [6], the standard elliptic curve over GF(2^m) is given by, $y^2+xy = x^3+ax^2+b$, where x, y $\in$ GF(2^m) , and b $\neq$ 0. Similarly the standard elliptic curve over GF(p) is

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

given by $y^2=x^3+ax^2+b$, where x, y $\in$ GF(p) and $4a^3+27b^2 \neq 0$ (mod p). For the ECC schemes, the most time-critical operation is the Elliptic curve scalar multiplication(ECSM), which consists of Elliptic curve Point addition(ECPA) and Elliptic curve point doubling(ECPD).The ECSM can be achieved in several co-ordinate systems. Affine co-ordinate system needs inversion and the cost of inversion will be high. In order to minimize the number of finite field inversions, projective coordinates are used in this work. In this design, we use Lopez's dahab Mixed coordinates[27] for ECPA and pure projective coordinates for ECPD over GF(2^m). For GF(p), Jacobian's Mixed coordinates for ECPA and pure projective coordinates for ECPD are used [16]. In this paper, we have used Montgomery ladder algorithm(Double and Add method), as it is suitable for our DF-ECC processor[33].

III. CARRY SAVE ADDER

The simple and straightforward method of adding n numbers (m bit wide) is to add the first two and then sum to the next and so on. But this method requires total gate delay of O(n lg m). So, by using carry save addition method delay can be reduced. The carry save addition approach is quite different from normal addition.

The carry save approach divides the process into three steps. The first step is to consider only the sum. The second step is to find the carry. The third step is to add this carry to the already calculated sum.

For example, if we want to add three numbers a, b and c, then the result is calculated in the form of sum and carry. The sum is equal to a+b+c mod 10.

Example:

```
a:   1 2 3 4 5
b:   1 2 3 5 2
c: +   2 1 3 4 5
-----------------
sum:   4 5 9 3 2
```

Here sum S_i equals the sum of $a_i+b_i+c_i$ modulo 10

```
a:   1 2 3 4 5
b:   1 2 3 5 2
c: + 2 1 3 4 5
-----------------
carry:  0 0 0 1 1
```

Carry is equal to the sum of the bits from the previous column divided by 10.The total is obtained by adding the sum with the carry as shown below.

```
Sum:   4 5 9 3 2
Carry: 0 0 1 1
---------------
       4 6 0 4 2
```

In summary, CSA can add several numbers at a time rather than just two. CSA is very useful in calculating the partial products in multiplication. CSA produces all of the output bits in parallel in the same delay as full adder(FA).

IV. VEDIC MATHEMATICS

Vedic Mathematics is the ancient system of mathematics but again which was rediscovered by Shri Bharati Krishna Tirthaji [30]. The word "Veda" is a Sanskrit word means "knowledge". Vedic Mathematics is a system of mental mathematics. Shri Bharati Krishna Tirthaji, produced a reconstruction of the ancient mathematical system based on sixteen sutras, together with a number of sub-sutras only after lengthy and careful investigation, Swami Bharati Krishna Tirtha (1884-1960), former Jagadguru Sankaracharya of Puri culled a set of 16 Sutras (aphorisms) and 13 Sub - Sutras (corollaries) from the Atharva Veda. He developed methods and techniques for amplifying the principles contained in the aphorisms and their corollaries, and called it Vedic Mathematics. The Sutras apply to and cover almost every branch of Mathematics. They apply even to complex problems involving a large number of mathematical operations [31]. Application of the Sutras improves the computational skills of the learners in a wide area of problems, ensuring both speed and accuracy, strictly based on rational and logical reasoning.

The list so compiled contains Sixteen Sutras are as stated hereunder.

1. EkadhikenaPurvena:-"By one more than the previous one".
2. NikhilamnavatascaramamDasatah:- "all from 9 and the last from 10".
3. Urdhva – tiryagbhyam: It is the general formula applicable to all cases of multiplication and also in the division of a large number by another large number. Just with two simple words is "Vertically and crosswise".
4. ParavartyaYojayet: -'transpose and apply'
5. SunyamSamyaSamuccaye: -'Samuccaya is the same, that Samuccaya is Zero.'
6. Anurupye – Sunyamanyat:- 'If one is in ratio, the other one is zero'.
7. Sankalana – Vyavakalanabhyam:-'by addition and by subtraction'.
8. Puranapuranabhyam:-"by the completion or non – completion in solving the roots for general form of quadratic equation.
9. Calana – Kalanabhyam: 'Sequential motion'.
10. YavadunamTavadunikrtyaVargancaYojayet:-'whatever the deficiency subtract that deficit from the number and write alongside the square of that deficit'.
11. Vyastisamastih—"Specific and General".
12. S`esanyankena Caramena—"The Remainders by the Last Digit".
13. Sopantyadvayamantyam—"The Ultimate and Twice the Penultimate".
14. EkanyunenaPurvena: This sutra comes as a Sub-sutra to Nikhilam which gives the meaning 'One less than the previous' or 'One less than the one before'.
15. GunitaSamuccayah -- is intended for the purpose of verifying the correctness of obtained answers in multiplications, divisions and factorizations.
16. Gunakasamuccayah – "All the multipliers"

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

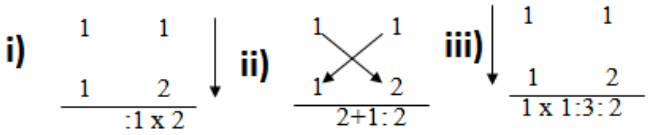
1. Urdhva – tiryagbhyam

This sutra, as the title suggests, translates to "Vertically and crosswise". This sutra is one of the best known of the Vedic Sutras, and has found many applications [28-29]. It is a general formula applicable in all cases of multiplication and also in division of a large number by another large number. Urdhva-triyagbhyam multiplier architecture shows speed improvements over conventional multiplier architecture[32]. An example is as shown below in three steps. The product of 11 X 12 can be determined as follows

i) The right hand most digit of the multiplicand, the first number (11) i.e.,1 is multiplied by the right hand most digit of the multiplier that is, the second number of (12) i.e., 2. The product 1 X 2 = 2 forms the right hand most part of the answer.

ii) Now, diagonally multiply the first digit of the multiplicand (11) i.e., 1 and the second digit of the multiplier (12) i.e., 1 (answer 1 X 1=1); then multiply the second digit of the multiplicand i.e.,1 and first digit of the multiplier i.e., 2 (answer 1 X 2 = 2); add these two i.e.,1+ 2 = 3. It gives the second digit of the answer, 3.

iii) Now, multiply the left hand digit of the multiplicand i.e., 1 and left hand digit of the multiplier i.e., 1 vertically, i.e., 1 X 1 = 1. It gives the left hand most part of the answer.
Thus the answer is 13 2.



This gives output as 132

V. PROPOSED MULTIPLIER ARCHITECTURES

Dual Field Multiplier: is implemented using two methods as follows.

A. Method 1 (M1)- Multiplier using CSA:

The design of 8x8 CSA multiplier architecture using method M1 is shown in Fig.1. This is designed using HAs, FAs and AND gates. Initially there is no previous carry, for the first stage HA is used, for the remaining stages FA is used. The output of this multiplier will be 16-bits. Control signal (CS) is used to select either GF(2^m) or GF(p).

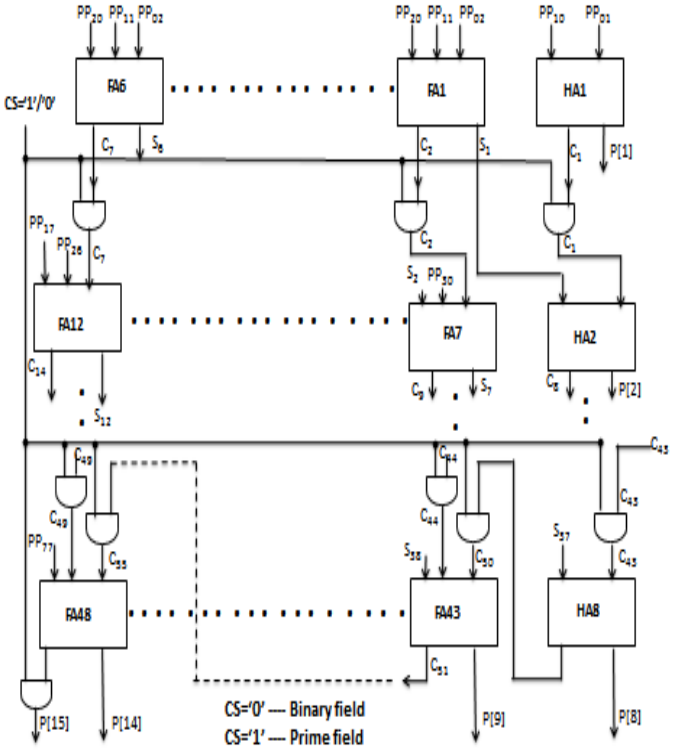


Fig. 1. Architecture of a 8x8 dual field multiplier using CSA.

B. Method 2 (M2)-Vedic Multiplier:

The design of 8x8 Vedic multiplier architecture using Method M2 which includes Urdhva – tiryagbhyam sutra and adder is shown in Fig.2. The inputs are a(a7-a0) and b(b7-b0). The output of this multiplier will be 16-bits as the carry C_{out} and the sum S_{14} to S_0.

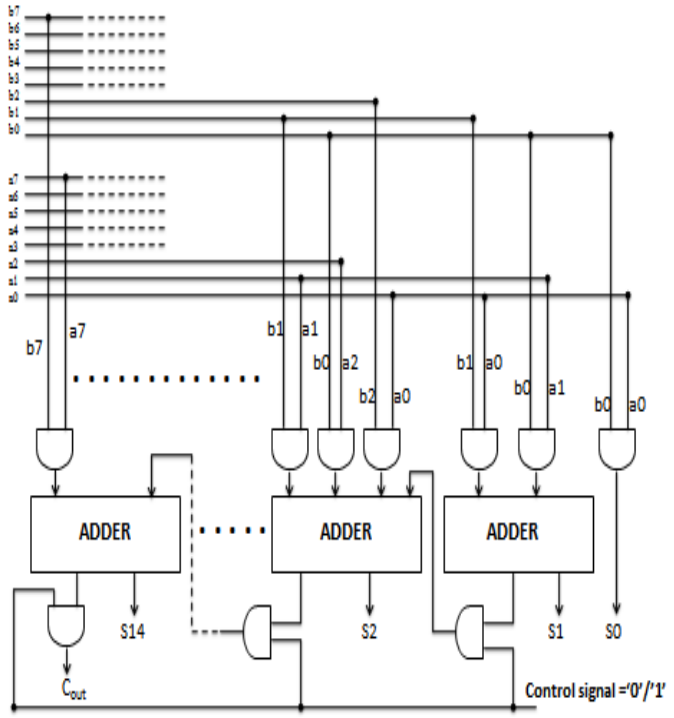


Fig. 2. Architecture of a 8x8 dual field Vedic multiplier

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

C. General architecture for method M1 or M2:

Fig.3. shows the general 16x16 multiplier using method M1 or M2. It is implemented using four 8x8 CSA multiplier or Vedic Urdhva-Tiryagbhyam multiplier modules, and two 16 bit Full adders (FA). The output of this multiplier will be 32 bits. That is, LSB of first 8bit multiplier [7:0], LSB of first FA sum output [7:0], sum output [14:0] of the second FA and carry. Thus finally we get 32bit product term. The field GF(2^m) or GF(p) is selected based on the value '1'or '0' of the Control signal (CS). If CS = '0', all the carry bits are disabled (made zero's), the output belongs to GF(2^m). If CS='1' the carry bits are enabled and the output belongs to GF(p).

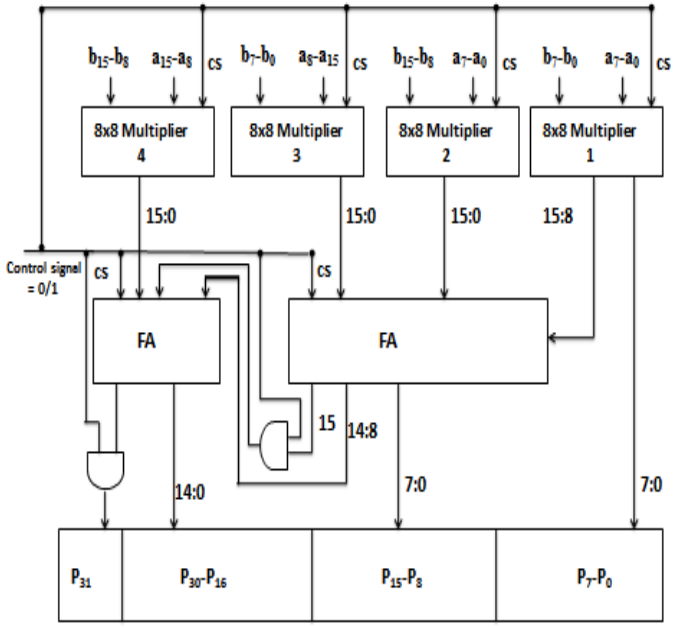


Fig. 3. General Architecture of a 16x16 multiplier

D. Proposed dual field adder:

The Dual Field (DF) adder is constructed using look ahead carry technique with the control signal is shown in Fig.4. If the control signal is held at '1', the DF adder acts as a prime field adder else (control signal ='0') the DF adder acts as a binary field adder. The inputs to the MOD (modulo p) block are final carry C_{out}, buffer output 'S' and p. The output of the MOD block gives the prime field adder output, that is serves as the first input to the 2:1 Multiplexer and the same buffer output 'S' gives the binary field adder output and serves as the second input to the 2:1 multiplexer. 2:1 multiplexer select signal is connected to the control signal. When the CS='0', the 2:1 multiplexer gives the binary field adder output else (CS='1') gives the prime field adder output. As shown in Fig.5, Cg and Cp are the carry generate and carry propagate signals. Here Cg= a.b and Cp= a + b as shown in Fig.5.

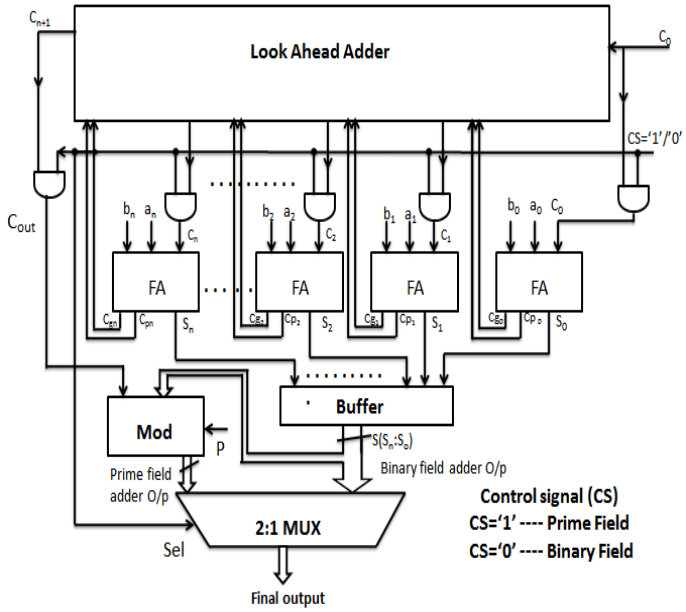


Fig. 4.DF- Adder using Look ahead adder technique

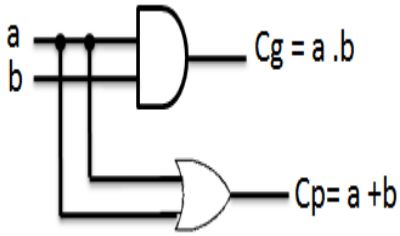


Fig.5.Cg and Cp block

VI. DUAL FIELD ECC PROCESSOR

The Dual field ECC processor is shown in Fig.6. The Dual field ECC processor block supports the following functions. 1. Linear Feedback Shift Register (LFSR), 2. ECC arithmetic unit and 3. Inversion Unit. Random numbers are generated using LFSR to improve security. Generations of different scaling factor 'k' values using LFSR are stored in RAM locations. To get kP, 'k' value is selected based on the address value given in the RAM locations. Scalar multiplication (kP) is done in ECC arithmetic unit. 'kP' is achieved using point doubling and point addition. Depending on the Individual bits of 'k' value, either point doubling or point addition is selected. Montgomery ladder algorithm (Double and Add method) is achieved using recurrent point addition and doublings. To perform Point doubling and addition, finite-field arithmetic operations are required. For example, add, multiplication, square, etc.. In Section 3 we have already discussed two multiplier architectures and an adder. These can be applied to either prime filed or binary field. ECC arithmetic unit provides the results as projective coordinates which are then given to the inversion unit. The function of the inversion unit is to generate the affine points using Extended Euclidean Inverse algorithm. These affine points are fed to any application.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

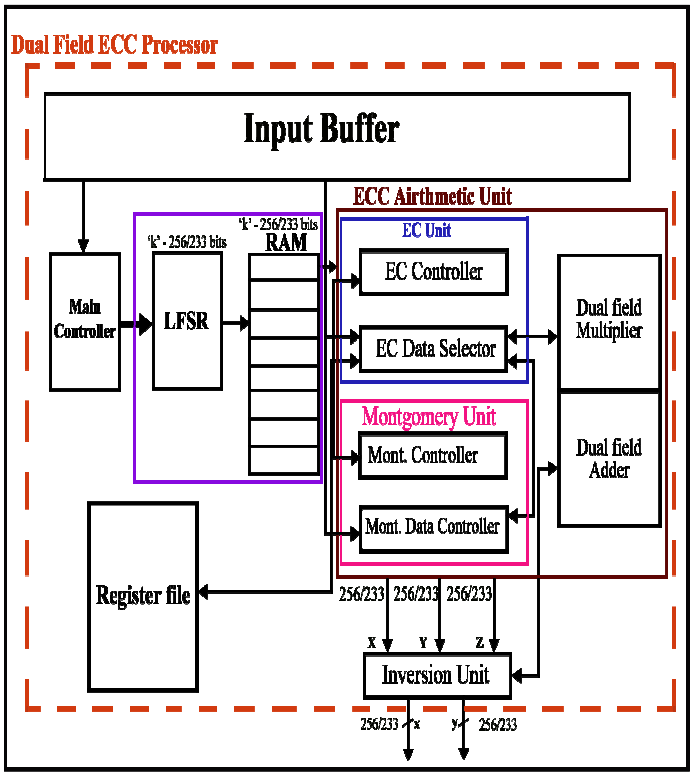


Fig. 6. Dual Field ECC processor

VII. EXPERIMENTAL RESULTS AND COMPARISONS

The experiment was conducted for the DF-ECC processor using the appropriate multipliers of size 256/233 bits on an XC5vlx110t-1ff1136 FPGA verification platform to evaluate area, speed and power throughputs. The total On-Chip Power is calculated by using Xilinx X-Power analyzer, which consists of Dynamic power and Quiescent power.

A. *Results for two different multiplier methods M1 and M2*

Synthesis results are shown in Table I. The results illustrate maximum operating frequency, number of slices occupied and the power consumed. From these results, we come to the conclusion that M2 multiplier is faster compared to M1.

B. *Implementation of Individual Fields*

We have described the two individual field ($GF(2^m)$) and GF(p)) using two different multiplier methods (M1 and M2). Table II shows the $GF(2^m)$ can run at **142.146 MHz** and **51.784MHz** over GF(p) using M1, and the same $GF(2^m)$ can run at **153.416MHz** and **56.439MHz** over GF(p) using M2. From these outcomes, we can conclude that, $GF(2^m)$ has a shortest critical path and GF(p) has a longer critical path.

C. *Implementation of DF-ECC processor*

We have described DF-ECC processor of 256/233 bits on an XC5vlx110t-1ff1136 FPGA verification platform to evaluate area, speed and power throughputs. Table III shows the synthesis results. Using M1, Our DF-ECC processor runs with a maximum frequency of **139.879MHz**, occupies 4179 slices, number of BRAM is 28 and consumes 1090mW of power. The same DF-ECC processor using M2 runs with a maximum frequency of **151.596MHz**, occupies 6463 slices, number of BRAM is 28 and consumes 1096mW of power.

The emulation result is shown in Fig. 7 and 8 and is verified by using chip scope-pro. The four output signals are shown, are known as Trig0 (prime1), Trig1 (prime2), Trig2 (binary1) and Trig3 (binary2). Here, prime1 and prime2 indicate prime field affine values and binary1 and binary2 indicate the binary field affine values. The input signals are required to see the emulation output waveform, and they are pbarb, clk, wr_en, lrst, addr. Signal 'addr' is the address location selected as '100' in this waveform. The signal 'pbarb' selects the prime field or binary field. If 'pbarb' is '0' selects the prime field is shown in Fig.7.Similarly 'pbarb' is '1' selects the binary field is shown in Fig.8.

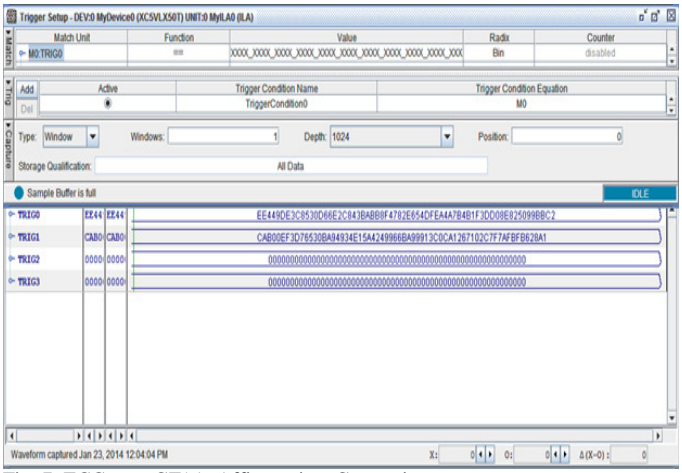


Fig. 7. ECC over GF(p). Affine points Generation

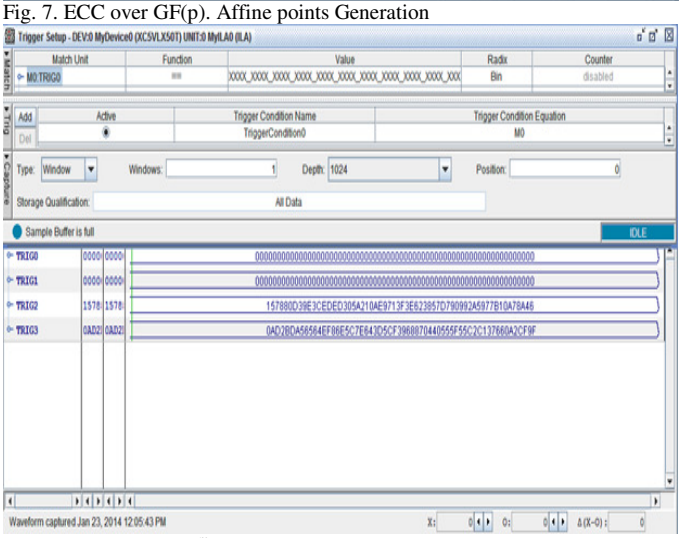


Fig. 8. ECC over $GF(2^m)$. Affine points Generation

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE I. TWO DIFFERENT MULTIPLIER DESINGS (M1 AND M2)

	Platform	*No. of Bits*	*f_{MAX} in MHZ*	*Area [Slices]*	*Power (W)*
Ours (M1)	**Xilinx Virtex-5**	**256**	**146.67**	**1596**	**1.066**
Ours (M2)	**Xilinx Virtex-5**	**256**	**160.45**	**2989**	**1.071**

TABLE II. COMPARISION AMONG THE EXSISTING ECC DESIGNS

	Platform	*Field Size*	*f_{MAX} in MHZ*	*Area [Slices]*	*Power (W)*
Ours (M1)	**Xilinx Virtex-5**	**233-bit GF(2^m)**	**142.146**	**2913**	**1.071W**
		256-bit GF(P)	**51.784**	**3427**	**1.085W**
Ours (M2)	**Xilinx Virtex-5**	**233-bit GF(2^m)**	**153.416**	**5102**	**1.084W**
		256-bit GF(P)	**56.439**	**5748**	**1.092W**
[9]	Xilinx Virtex-2 pro	256-bit GF(P)	43.51	2662 CLB slices	-----
[22]	Xilinx Virtex-2 pro	256-bit GF(P)	37.037	8272 CLB slices	----
[28]	Xilinx Virtex-4	233-bit GF(2^m)	142.53	2648	----
[26]	Xilinx Virtex-5	163-bit GF(2^m)	147	3513	1.078W
[23]	Xilinx XC4vLX 80	163-bit GF(2^m)	143	24,363	----

TABLE III. COMPARISION AMONG THE EXSISTING ECC PROCESSOR DESIGNS USING METHOD 1& 2 MULTIPLIER.

	Platform	*ECC processor field size*	*f_{MAX} in MHZ*	*Area [Slices]+ BRAM's*	*Dual Field*	*Power (W)*
Ours (M1)	**Xilinx Virtex -5**	**233/256**	**139.879**	**4179+ 28**	**Yes**	**1.090W**
Ours (M2)	**Xilinx Virtex -5**	**233/256**	**151.596**	**6463+ 28**	**Yes**	**1.096W**
[24]	Xilinx Virtex-2 pro	160/160	100	8,594 +6	Yes	--------
[5]	Xilinx Virtex-2 pro	160/256	94.7	39,531 / 41,595	Yes	---------
[25]	Xilinx XC4Vfx100	163/192	150.5	5,227 CLB's	Yes	---------

D. Comparisons

Table II compares existing ECC designs in two different fields separately. And Table III compares our DF- ECC processor when scaled up to 256/233 bits ECC. In Table III the two methods have been incorporated. Comparing the Method 1 and Method 2, our DF-ECC processor using Method 2 multiplier design is more efficient in terms of speed and power.

VIII. CONCLUSION

We have proposed two new multiplier architectures for DF-ECC processor. ECC arithmetic unit requires the finite field arithmetic operations. It has been implemented using M1 and M2 on a Virtex 5 Xilinx FPGA device.

Using M1 multiplier design, with a prime field of 256 bits the point scalar multiplication with coordinate conversion operates at a maximum frequency of 51.784MHZ, 1085mW power and occupies 3427 slices over a GF(p). Also, using M1,the point scalar multiplication with coordinate conversion operates at a maximum frequency of 142.146MHZ, 1071mW power and occupies 2913 slices over a GF(2^m). Similarly using M2 multiplier design, the point scalar multiplication with coordinate conversion operates at a maximum frequency 56.439MHZ, 1092mW power and occupies 5748 slices over a GF(p). Also, using M2, the point scalar multiplication with coordinate conversion operates at a maximum frequency of 153.416MHZ, 1084mW power and occupies 5102 slices over a GF(2^m). The experimental results prove that our DF-ECC processor architecture using M2 runs with a maximum frequency of 151.596MHz occupies 6463 slices and consumes 1096mW of power. Also, our DF-ECC processor provides comparable improvements in terms speed and power with Existing DF-ECC processor designs.

In future, Hardware realization of high speed elliptic curve point multiplication using precomputation over GF(p) can be investigated. The existing investigation can be extended to address varied design parameters like speed, power and area.

ACKNOWLEDGMENT

The author would like to thank RNS Institute of Technology for the constant support and Encouragement.

REFERENCES

[1] N. Koblitz, "Elliptic curve cryptosystems," Math. Comput., vol. 48,pp. 203–209, 1987.

[2] V.S. Miller, "Use of elliptic curve in cryptography," in Proc. Crypto,1986, pp. 417–426.

[3] B. Schneier, Applied Cryptography: Protocols, Algorithms, and Source Code in C, 2nd ed. New York: Wiley, 1996.

[4] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, Handbook of Applied Cryptography. Boca Raton, FL: CRC Press, 1997.

[5] IEEE 1363 standard specifications for public-key cryptography,1363,Jan. 2000.

[6] NIST, Recommended Elliptic Curves for Federal Government Use, May 1999 (http://csrc.nist.gov/encryption).

[7] William N. Chelton, "Fast elliptic curve cryptography on FPGA," IEEE Transactions on VLSI, Vol. 16, No. 2, February 2008.

[8] Henri Cohern, "Efficient Elliptic Curve Exponentiation Using Mixed Coordinates".

[9] Chang Hoon Kim , Soonhak Kwon , Chun Pyo Hong , "FPGA implementation of high performance elliptic curve cryptographic processor over GF(2^{163}) Journal of Sys. Architecture,Elsvier,2008.

[10] Jyu-Yuan Lai and Chih-Tsun Huang, "A Highly Efficient Cipher Processor for Dual-Field Elliptic Curve Cryptography," IEEE Transactions on Circuits and Systems—II,vol. 56, No. 5, May 2009

[11] Prabhat Chandra Shrivastava, Rupesh Kumar, ArvindKumar,and Sanjeev Rai "High-Speed and Low Power Unified Dual-Field Multiplier in GF (P) and GF (2^m)," IEEE, 2010.

[12] Yi Wang, Douglas L. Maskell, Jussipekka Leiwo, "A unified architecture for a public key cryptographic coprocessor," Journal of systems Architecture, Elsvier,2008.

[13] Sujoy Sinha Roy, Chester Rebeiro, and Debdeep Mukhopadhyay, "Theoretical Modeling of Elliptic Curve Scalar Multiplier on LUT-Based FPGAs for Area and Speed", IEEE Transactions on VLSI systems, Vol. 21, No. 5, May 2013

[14] Antonio Cortina Reyes ,Ana Karina Vega Castillo, Miguel Morales-Sandoval and Arturo D´ıaz-P´erez, "A Performance Comparison of Elliptic Curve Scalar Multiplication Algorithms on Smartphones",IEEE, 2013

[15] I. Blake, G. Seroussi, N. Smart, Elliptic Curves in Cryptography,London Mathematical Society Lecture Note Series 265, Cambridge University Press, 2000

[16] D. Hankerson, A. Menezes, S. Vanstone, Guide to Elliptic Curve Cryptography, Spinger, United States,2003

[17] Akashi Satoh and Kohji Takano, "A Scalable Dual-Field Elliptic Curve Cryptographic Processor," IEEE Transaction on Computers, vol. 52, No. 4,Apr 2003.

[18] Jyu-Yuan Lai and Chih-Tsun Huang, "High-Throughput Cost-Effective Dual-Field Processors and the Design Framework for Elliptic Curve Cryptography," IEEE Transactions on VLSI systems, vol. 16, No 11,Nov 2008.

[19] Wanzhong Sun, Zibin Dai and Nianmin Ren2., "A Unified, Scalable Dual-Field Montgomery Multiplier architecture for ECCs," IEEE,2008.

[20] B.MuthuKumar, S.Jeevananthan,"High Speed Hardware Implementation of an Elliptic Curve Cryptography(ECC)Co-Processor," IEEE 2010

[21] Jen-Wei Lee, Ju-Hung Hsiao, Hsie-Chia Chang, and Chen-Yi Lee, "An Efficient DPA Countermeasure With Randomized Montgomery Operations for DF-ECC Processor", IEEE Transactions on Circuits and Systems—II,vol. 59, No. 5, May 2012.

[22] Gustavo D. Sutter, Jean-Pierre Deschamps, and José Luis Imaña, "Efficient Elliptic Curve Point Multiplication Using Digit-serial Binary Field Operations," IEEE Trans. On Industrial Elns, Vol. 60, No. 1, Jan 2013.

[23] K.C.Cinnati Loi and Seok-Bum Ko, "High Performance Scalable Elliptic Curve Cryptosystem Processor in GF(2^m),IEEE,2013.

[24] Jen-Wei Lee, Szu-Chi Chung, Hsie-Chia Chang and Chen-Yi Lee, "Efficient Power-Analysis-Resistant Dual-Field Elliptic Curve Cryptographic Processor Using Heterogeneous Dual-Processing Element Architecture," IEEE Trans. On VLSI systems,2013.

[25] K Sakiyama, L. Batina, B. Preneel, and I. Verbauwhede, "Multi-corecurve-based cryptoprocessor with reconfiguarable modular. arithmetic logic units over GF(2^m)" IEEE Trans. On Computers, vol. 56, no. 9, pp. 1269–1282, Sep. 2007.

[26] K. Sakiyama, E. D. Mulder, B. Preneel, and I. Verbauwhede, "A Parallel processing hardware architecture for elliptic curve cryptosystems," inProc. IEEE ICASSP, Toulouse, France, May 2006, vol. 3, pp. 904–907.

[27] J. Lopez and R. Dahab, "Improved algorithms for elliptic curve arithmetic in GF(2^m)" in Proc. Sel. Areas Cryptography: 5th Annu. Int.Workshop (SAC), vol. 1556, pp. 201‑212 ,Aug. 1998.

[28] Thapliyal H. and Srinivas M.B. "High Speed Efficient N x N Bit Parallel Hierarchical Overlay Multiplier Architecture Based on Ancient Indian Vedic Mathematics", Transactions on Engineering, Computing and Technology, Vol.2, 2004.

[29] N. Shylashree, D. Venkata Narayana Reddy, V. Sridhar, "Efficient implementation of RSA encryption and decryption using Ancient Indian Vedic Mathematics", CiiTInternational journal of Programmable Devices Circuits and Systems" June 2012, India, and Print: ISSN0974-973X & online: ISSN 0974-9624.

[30] Jagadguru Swami Sri Bharath, Krishna Tirathji, "Vedic Mathematics or Sixteen Simple Sutras from The Vedas", Motilal Banarsidas, Varanasi (India), 1992.

[31] Jeganathan Sriskandarajah, "Secrets of Ancient Maths: VedicMathematics", Journal of Indic Studies Foundation, California, 2003 pp. 15-16

[32] Parth Mehta, Dhanashri Gawali, "Conventional versus Vedicmathematical method for Hardware implementation of a multiplier" ; in Proc. IEEE Int. Conf. Advances in Computing, Control, and Telecommunication Technologies, India, Dec. 2009, pp. 640-642.

[33] N.Shylashree and V. sridhar "A Unified Architecture for a Dual Field ECC Processor Applicable to AES" Fifth International conference on CIMSim 2013, IEEE,2013.

AUTHORS PROFILE

N.Shylashree received B.E. degree in Electronics and Communication from VTU, Karnataka, India in 2006. She received M.Tech. degree in VLSI Design and Embedded Systems from VTU, Karnataka, India in 2008. She is currently a research scholar(part-time) in PESCE, Mandya, Karnataka, India and is also an Assistant Professor in ECE Department, RNSIT, Karnataka, India. She has published eight research papers in various international journals and conferences

Dr. V. Sridhar is currently The Principal of P.E.S. College of Engineering, Mandya, Karnataka, India. He was the Registrar(Evaluation) in VTU, India. His research interests include Cryptography, VLSI and Embedded Systems, and Bio-Medical Engineering. He has published more than 50 research papers in various international journals and conferences

Wavelet based Iris Recognition for Biometric Systems

Jitendra Singh Uriaya
Department of Electronics and Communication
Jabalpur Engineering College,
Jabalpur India (482011) .

Agya Mishra
Department of Electronics and Communication
Jabalpur Engineering College,
Jabalpur India (482011) .

Abstract— **Iris recognition is a critical area of research in the field of security system and personal identification. In this paper, a novel, efficient technique for iris recognition is presented. The aim is to develop a lifting wavelet based algorithm. This method reduces the noise to the maximum extent possible, and extracts important information from the image. In this paper recognition is being done biorthogonal wavelet based method which is then compared with other wavelet family. The method was tested on the CASIA dataset of iris image. For matching the iris template Hamming Distance technique is used and checks the accuracy of the system. The present algorithm provides the 94.4% recognition rate and accuracy approx. 99 % on the CASIA iris data base version 1.**

Keywords*— *Iris recognition, Biorthogonal wavelet, Haar wavelet, Daubechies wavelet, Iris template, Hamming distance, Match score.

I. INTRODUCTION

With more efficient communication techniques and more complex network societies, human recognition has gained significance. Biometrics is defined as automated methods for verifying or identifying the identity of a living individual based on physiological or behavioral characteristics [1]. Biometrics has advantages over traditional identifiers like identity cards, signatures etc, in that they cannot be forgotten, transferred, misplaced, duplicated, or stolen easily [1], [2]. Some examples of biometric identifiers include fingerprints, iris, hand geometry, voice, speech, face and gait. In generally biometric identification is gaining more popularity and more acceptances in public as well as in private sectors. Iris recognition is considered to be highly accurate and reliable method of biometric identification. The iris being found to be very stable, highly unique and easy to capture, is classified as one of the better biometric identifiers [3, 4]. The unique epigenetic pattern of a human iris is used for personal identification. Image processing and signal processing techniques are employed to extract information from unique iris structure from a digitized image of an eye [5, 6, 7, 8]. This information is encoded to formulate a "biometric template", which is stored in a database and also used for identification. Thus, the purpose of the template formation is to mathematically encode the iris pattern and match it with other similar representations. To perform iris recognition, an algorithm detects the iris from eye image and performs the segmentation, normalization and create iris template. There are several categories of iris recognition algorithms, such as Daugman's algorithm, curvelet transform [9], imaginary coefficient of morlet wavelet transform[10], novel biorthogonal wavelet[11], combined feature of GLCM and wavelet transform[12], hypercomplex Gabor wavelet[13] and dual tree complex wavelet transform[14]. Among the data-driven methods, Daugman's algorithm is one of the most popular algorithms [4, 15, 16, and 17]. The process of iris recognition is divided in four steps namely, segmentation where the iris region is isolated in an image under consideration; normalization where the pixel of the isolated iris is mapped from concentric domain to non-concentric domain; feature extraction it provides abundant texture information, a feature vector is formed which consist of the order of sequence of feature extracted from various representation of the iris image; matching where 'Hamming distance' between each pair is calculated.

In this paper, wavelet based iris recognition is being done by using biothogonal wavelet which is then compared with other wavelet family and experimental results are taken for measuring hamming distance and to check the accuracy of the system. To understand the basic concept of iris recognition let us discuss in next section. This paper is organized as follows. Section II wavelet based iris recognition. Section III presents the experimental analysis Section IV comparison table with measuring parameter. Section V finally concludes the paper.

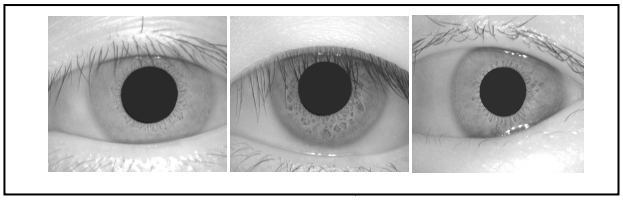


Fig.1 Sample of iris image

A. Data Management

It is important to test the designed algorithm on sufficiently large and diverse data set. A data set of gray scale iris digital images taken from the Chinese Academy of Science (CASIA version 1) is used for testing. The data base consists of 756 gray scale images coming out of 108 distinct classes and 7 images of each eye. Sample of iris image is shown in fig.1. It was concluded by inspecting visually that the iris image have sufficient resolution and clarity and hence no further

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

enhancement technique are used. The available data can be tabulated as follows.

TABLE1. DATA SET DISTRIBUTION

Data Set	No. of images	No. of Classes	Resolution
CASIA	756	108	320*280

B. Wavelet Analysis

For the study of wavelet analysis, wavelet is used in the digital images to formulate a template. Replacing the traditional multiresolution analysis (MRA) scheme, a lifting technique is used to construct the biorthogonal filters [17, 18, 19, 20]. The main advantage of this scheme over the classical construction is that it does not based on the Fourier transform. Also, it allows faster implementation of wavelet transform. The basic idea behind the lifting scheme is shown in Fig. 2. It starts with trivial wavelet, the ``Lazy wavelet"; which has the formal properties of wavelet, but is not capable of doing the analysis. The lifting scheme then gradually builds a new wavelet, with improved properties, by adding in a new basis function. This is the inspiration behind the name of the scheme. The lifting scheme can be visualized as an extension of the FIR (Finite Impulse Response) schemes [21].

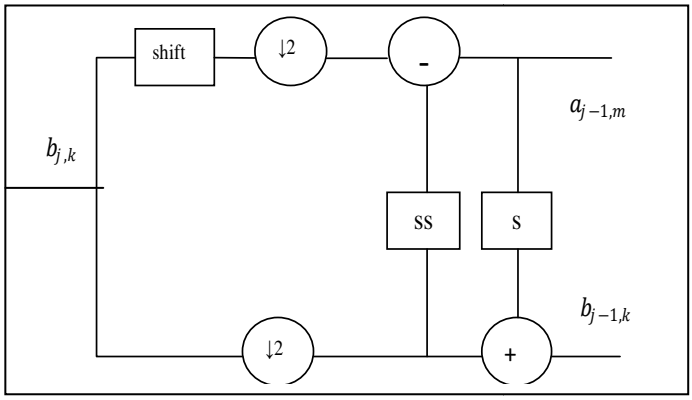


Fig 2. The lifting scheme for wavelet. It first calculates the Lazy wavelet transform, then calculate the $a_{j-1,m}$ and finally lifts the $b_{j-1,k}$.

It is known that any two channel FIR sub band transform can be factored into finite sequence of lifting steps is faster and efficient.

II. WAVELET BASED IRIS RECOGNITION

This portion describes the algorithm development to perform iris recognition for ideal condition. The algorithm is flexible enough to be adapted for the non-ideal cases like off angle images, noise etc. The algorithm can be divided in to four steps which are as:

- Segmentation
- Normalization
- Template Formation or Encoding
- Match score calculation

The iris recognition consists of the following four steps. The block diagram of proposed iris recognition algorithm system is shown in fig 3.

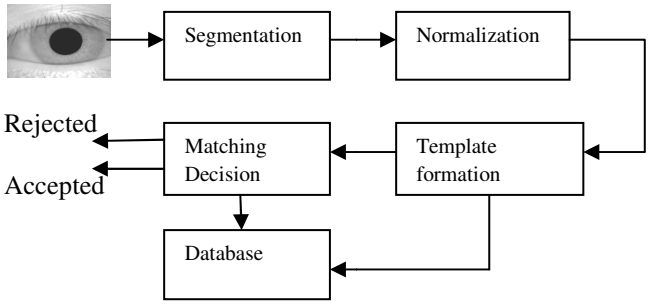


Fig 3. Block diagram of proposed algorithm iris recognition

C. Isolation and Segmentation

Segmentation involves isolating the important information from the rest of the eye image. Before doing actual segmentation, all the images are transformed into the wavelet domain and maxima energy extraction is performed. The number of retained coefficients is decided to be 10000. The images are inverse transformed before segmentation is performed using the log, simple gradient masking, and circular canny edge detection methodologies. The detected edges are mapped on the inverted iris images. The area within the global edge is phase shifted and then the complete image is transformed into the wavelet domain. The phase information is used and the marked area is further divided into smaller regions and average intensity thresholding is performed to further remove non-significant information. Non-significant information could be in the form of eye-lashes, eyelids and reflections [22, 23]. In order to take care of these different types of noise entities, thresholding is done for each sub-band and hence the scheme is called `in-band noise removal'. Each band significantly removes the noise content depending on whether it is high pass or low pass and the direction in which that filters acts i.e. horizontal, vertical or diagonal. E.g. eyelashes were predominantly removed in the HL band. The threshold value is adaptive and changes for every image depending upon the relative variations in the intensity of different parts of the image. The point where average intensity drops below the threshold is chosen as the iris radius. The threshold's effectiveness depends upon the angular resolution. However, increasing angular resolution may result in an excessively large template.

The circular resolution value of 20 is selected for this study, which is trade-off between providing effective division for noise removal as well as obtaining

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

reasonable size templates. The isolated iris images are shown in fig 3.

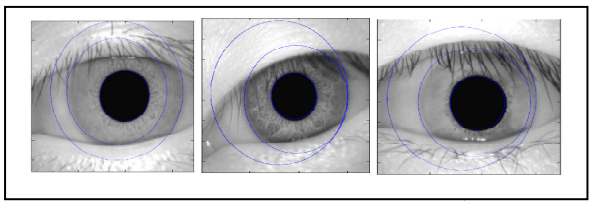


Fig .3 Isolated iris region. The outer and inner boundaries of the iris region are marked with circle.

D. *Normalization*

In this section our algorithm involves the normalization process. In the normalization process first map the data from (r, θ) domain to (x, y) domain. For this the centre of iris is taken as the starting point as against pupil centre in the Daugman's method. The linear radial vector traces the complete iris region. Also, instead of matching up the iris and pupil centres, the vector dimension of the vector is kept fixed. This is the resolution along the iris radius. The vector dimension is fixed at *200*. For highly non-concentric iris pupil pairs matching of the centres is performed. The diagonal information is mapped in the opposite quadrant to avoid any loss of information. The normalized iris shown in fig 4.

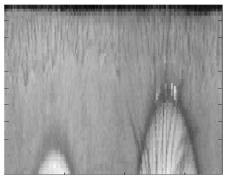


Fig 4. Normalized iris

D. *Template formation and Encoding*

In this section, iris template is formed. The template formation use the segmented and normalized iris information is transformed into the wavelet domain using the biorthogonal tap. Lifting technique are used for designed the filter, which has the advantage that it is completely invertible. These filters transform the data into a different and new basis where 'large' coefficient corresponds to relevant image data and 'small' coefficient correspond to the noise. Thresholding is performed once again. The process is known as image de-noising. Wavelet encoded data is scalable and localized and hence matching of the feature at same location using various scales is possible. A haar wavelet and daubchies wavelet are used for comparison. On these wavelet lifting technique is also performed. For the creation of the iris template we perform the train process. In this process we take total iris classes 10 and image per class 3. This is done for all the iris images and forms the iris template. Along with the iris template, the mask template is also formed which is used for locating the noisy parts of the image.

E. *Matching Score Calculation*

In this process only the phase information is used and converted to ones and zeros are encoded from the normalized iris pattern. The formulated template is taken for matching and similarity scores are calculated. Bit wise comparison of the template is made and hamming distance is calculated for every such comparison. This is obtained by doing successive bit wise "X-OR"ing and "AND"ing. To account for the rotational inconsistencies the maximum matched value is chosen. The mask templates are used to ignore the noisy parts of the image the formula for finding the hamming distance is given in [15] as,

$$\mathrm{HD} = \frac{\square(code\ A \otimes code\ B) \cap mask\ A\ \cap mask\ B\ \square}{\square\ mask\ A\ \cap mask\ B\ \square} \quad (1)$$

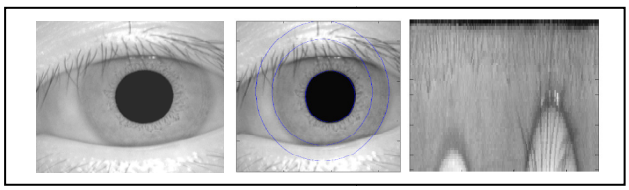
The "X-OR" ($\otimes$) operation detects the dissimilarity between the corresponding pair of bits, while "AND" ($\cap$) operation with mask makes sure that the noisy or less significant portion of the image is encoded. The ($\square\square$) represents the norm of the vectors. The Hamming Distance are normalized on the scale of a range from 0 to 1. The matching score are divided into inter-class and intra-class matching. The distributions are graphed based on these calculated values. Inter-class results are the results of matching iris of a person with iris of another person while intra-class results are the results by matching template of a person with another template of the same iris captured at other time. For calculating and separating the inter-class and intra-class distribution (*d*) Daugman[15] suggested a formula. If μ_1 and μ_2 are the means and σ_1 and σ_2 are the standard deviations of the distributions, then *d* is calculated as,

$$d = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 - \sigma_2^2)/2}} \quad (2)$$

III. EXPERIMENTAL ANALYSIS

Our algorithm has been performed on the gray scale iris image of bmp format and algorithm is tested on 756 iris images of the CASIA iris dataset. Canny edge detection is used to find out the edges of the iris and hough transform is used to detect circles in binary image. Experiment is performed on iris image by changing different wavelet from the wavelet family. Experimental results are shown along with iris and measuring parameter details.

Case Study I. Test iris image named database_2folder_2 has bmp format gray scale 320×280 and biorthogonal (bior3.5) wavelet is used. Figure no.4 shows the iris image, segmentation and normalization of iris. In fig. 4 the Hough transform detects circles with specific radius in a binary image and canny edge detection method is used to detect the edges of the iris.



http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Fig. 5 shows the results of biorthogonal wavelet filter. The hamming distance are plotted for biorthogonal filter. The intra class values treated as FRR values and inter-class values are treated as Far values and normalized occurrence rate are plotted on the y-axis as shown in the table below.

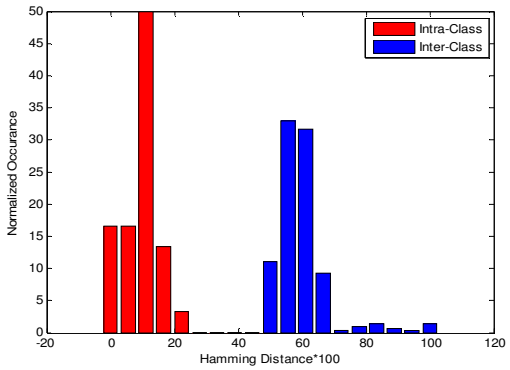


Fig. 5 Intra-class and inter-class Hamming distance distribution for biorthogonal filter.

Case Study II. In the case II test iris image named database_folder10_1 has bmp format gray scale 320×280 and haar wavelet is used. Figure no.6 shows the iris image, segmentation and normalization of iris. In fig. 6 the Hough transform detects circles with specific radius in a binary image and canny edge detection method is used to detect the edges of the iris.

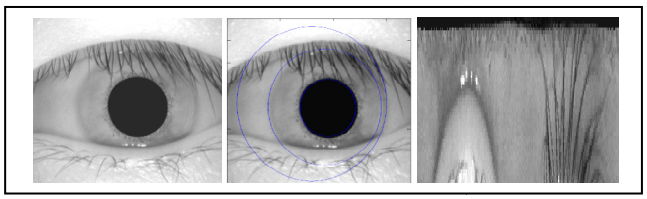


Fig 6. Sample of input iris, Segmented iris and Normalized iris

Fig. 7 shows the results of haar wavelet filter. The hamming distance is plotted for haar filter. The intra class values treated as FRR values and inter-class values are treated as Far values and normalized occurrence rate are plotted on the y-axis as shown below in table 2.

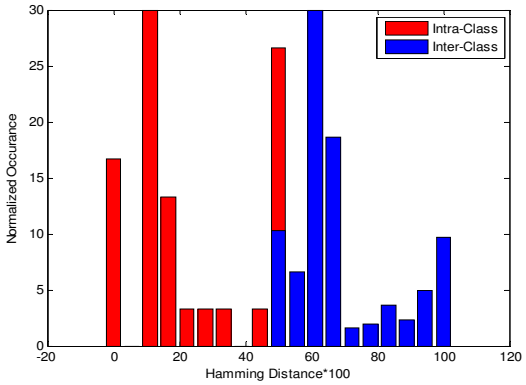


Fig. 7 Intra-class and inter-class Hamming distance distribution for Haar filters

Case Study III. In the case III test iris image named database_folder10_1 has bmp format gray scale 320×280 and daubchies wavelet is used. Figure no.8 shows the iris image, segmentation and normalization of iris. In fig. 8 the Hough transform detects circles with specific radius in a binary image and canny edge detection method is used to detect the edges of the iris.

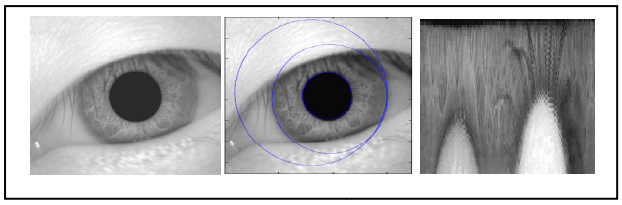


Fig 8. Sample input of iris, Segmented iris and Normalized iris

Fig. 9 shows the results of Daubechies wavelet filter. The hamming distance is plotted for daubechies filter. The intra class values treated as FRR values and inter-class values are treated as FAR values and normalized rate are plotted on the y-axis as shown below in table 2 .

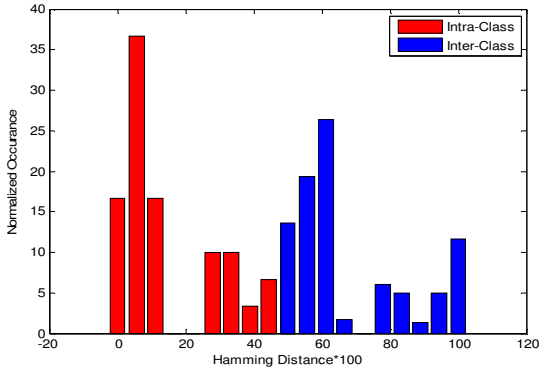


Fig. 9 Intra-class and inter-class Hamming distance distribution for Daubechies filter

IV. COMPERISON OF EXPERIMENTAL ANALYSIS

The present algorithm has been tested on 30 iris image of CASIA dataset and there are three different wavelets are used to check the accuracy of the algorithm by using three different iris images as shown above in experimental analysis. Fig 5, 7, 9 gives the results in detail. The hamming distance distribution is plotted for selected type of filter on the x-axis and normalized rate are plotted on the y-axis. The exact value of FRR (FALSE REJECTION RATE) and FAR (FALSE ACCEPTANCE RATE) are calculated by calculating the distribution of the hamming distance. By using the biorthogonal wavelet filter the normalized occurrence rate is much better as compared to by using haar and daubechies wavelet filter. Table shown below which show the comparison of different measuring parameters.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE 2 RESULTS COMPARISON OF FRR AND FAR FOR THE SELECTED WAVELET FILTER TYPE

Wavelet	FAR	FRR	Recognition rate %	Accuracy %
Biorthogonal	0.0	0.12	94.4	98.9
Daubchies	0.11	0.02	92.5	96.6
Haar	0.0	0.01	94	97

V. CONCLUSION

In this paper, a iris recognition method has been presented which is based on the wavelet. The presented algorithm can easily detect the iris and perform the iris recognition. The experimental results show that this method gives better recognition rate and accuracy by using biorthogonal wavelet as compared to the haar and daubchies wavelet. Our algorithm has large advantages in internet security and person identification as experimented in border monitoring and computer login. Although this method concentrate on accuracy of the systems. But the method can prove to be useful in person identification and for security purpose. Presented algorithm has been implemented on bmp image format. The results obtained show the accuracy of the presented method.

REFRENCES

[1] J. D. Woodward, N. M. Orlans, and P. T. Higgins, Biometrics. McGraw-Hill/Osborne, 2003.

[2] A. Jain, R. Bolle, and S. Pankanti, Biometrics: Personal Identification in Networked Society. Kluwer Academic Publisher, 1999.

[3] S.Noh, K.Pae, C.Lee, and J.Kim, Multiresolution independent component analysis for iris identification, The 2002 International Technical Conference on Circuits/Systems, Computer and Communication, Phuket,Thailand., 2002.

[4] R. Wildes, Iris recognition: an emerging biometric technology proceedingof the IEEE, vol.85, no.9, 1997.

[5] J. Daugman, .Biometric personal identification system based on iris analysis, United States Patent, vol. Patent Number: 5,291,560, 1994.

[6] W. Boles and B. Boashash, .A human identi_cation technique using images of the iris and wavelet transform,. IEEE Transactions on Signal Processing, vol. 46, no. 4, 1998.

[7] L. Ma, Y. Wang, and T. Tan, .Iris recognition using circular symmetric filters. national laboratory of pattern recognition,. Institute of Automation, Chinese Academy of Sciences, 2002.

[8] R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolczynski, J. Matey, and S. McBride, .A system for automated iris recognition,. Proceedings IEEE Workshop on Applications of Computer Vision, Sarasota, FL, pp. 121.128, 1994.

[9] Hanene Guesmi, Hanene Trichili, Adel M.Alimi, Basel Solaman "Iris verification system Based on Curvlet Transform" IEEE 2012.

[10] Zhonghua Lin, Bibo Lu "iris recognition method based on the Imaginary Coefficient of Morlet wavelet Transform" IEEE 2010.

[11] Aditya Abhyankar and Stephanie Schuckers "Novel Biorthogonal wavelet based iris recognition for Robust Biometric System" IACSIT April 2010 .

[12] V.V.Satyanarayana Tallapragada and E.G.Rajan "Iris recognition based on combined feature of GLCM and Wavelet transform" IEEE 2010.

[13] Fares S.Al-Qunaieer and Lahouari Ghouti," Color iris recognition using hypercomplex Gabor wavelet" IEEE 2009.

[14] A.S.Narote, S.P. Narote, L.M. WAglunare and M.B.Kokare"Robust iris feature extraction using dual tree complex transform" IEEE 2007.

[15] J. Daugman, .How iris recognition works,. Proceedings of 2002 International Conference on Image Processing, vol. 1, 2002.

[16] J. Daugman, 'Biometric decision landscapes', Technical Report No. TR482, University of Cambridge Computer Laboratory, 2002.

[17] W. Sweldens, 'The lifting scheme: A new philosophy in biorthogonal wavelet constructions,. in Wavelet Applications in Signal and Image Processing III, A. F. Laine and M. Unser, Eds. Proc. SPIE 2569, 1995, pp. 68.79.

[18] I. Daubechies and W. Sweldens, .Factoring wavelet transforms into lifting steps,. J. Fourier Anal. Appl., vol. 4, no. 3, pp. 245.267, 1998.

[19] R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, Wavelet transforms that map integers to integers,. Appl. Comput. Harmon. Anal., vol. 5, no. 3, pp. 332.369, 1998.

[20] W. Sweldens, .The lifting scheme: A construction of second generation wavelets,. SIAM J. Math. Anal., vol. 29, no. 2, pp. 511.546, 1997.

[21] D. Z.W. Kong, .Accurate iris segmentation based on novel reflection and eyelash detection model,. Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing, vol. Hong-Kong, 2001.

[22] N. Ritter, .Location of the pupil-iris border in slit-lamp images of the cornea,. Proceedings of the International Conference on Image Analysis and Processing, 1999.

Performance of Watermarking System using Wavelet Column Transform under Various Attacks

Dr. H. B. Kekre
Senior Professor
Computer Engineering Department
MPSTME, NMIMS University,
Vile Parle, Mumbai, India, .

Dr. Tanuja Sarode
Associate Professor
Computer Department,
Thadomal Shahani Engg. College,
Bandra, Mumbai 50, India

Shachi Natu
Ph.D. Research Scholar,
Computer Engineering Department
MPSTME, NMIMS University,
Vile Parle, Mumbai, India

***Abstract*— In this paper a novel wavelet transform based color image watermarking technique has been proposed. Column wise DCT wavelet transform has been applied to cover image. A watermark of size 128*128 is compressed and embedded in the cover image of size 256*256 using DCT wavelet column transform. Compression of watermark reduces the payload of information to be embedded in the cover image without significant loss of information. Robustness of proposed technique is tested against three types of attacks namely, cropping, Binary distributed run length noise and Gaussian distributed run length noise. Performance of proposed technique is observed for various compression ratios 1.45, 1.6, 1.77, 2, 2.285 and 2.67 beyond which the distortion in compressed watermark is perceptible. This technique proves to be exceptionally robust for Binary distributed run length noise attack. For Gaussian distributed run length noise attack also acceptable robustness is achieved. For cropping also it gives good performance except compression ratio 2. Higher compression of watermark results in better imperceptibility of watermarked image. Also higher compression ratio leads to increased robustness against attacks on watermarked image.**

***Index Terms*— Watermarking, DCT wavelet, Column transform, Binary distributed Run length noise, Gaussian distributed Run length noise.**

I. INTRODUCTION

With the rapid growth in use of internet technology and multimedia data for communication, importance of copyright protection of these digital data is also growing. Digital watermarking is one of such popular techniques for hiding information into digital data to protect it from unauthorized alterations. For invisible digital image watermarking, imperceptibility and robustness are of major concern and should go hand in hand though they are contradictory sometimes. Robustness refers to the ability to detect the watermark even after some processing of watermarked images is done. Imperceptibility means the existence of watermark in the digital image should not be noticeable to user. In literature, varieties of techniques have been implemented to address these problems.

In [1] Hu Guan, Zhi Zeng, Shuwu Zhang, have proposed a novel digital image watermarking algorithm, which embeds a multi-bit message into the image and performs the blind extraction of the message. To embed the watermark, the image is firstly scrambled to enhance the security. Then, some feature vectors based on the DCT of the scrambled image are extracted. Multi-bit watermark of length N is transformed into another form using spread spectrum technique. The expanded watermark generated using the spread spectrum technique is then embedded into the extracted feature vectors. Authors have also considered the Watson perceptual model to obtain good invisibility. By performing corresponding inverse transformations, watermarked image is obtained. The codebook in spread spectrum watermark generation is subjected to Gram-Schmidt's orthogonalization and unitization operation. This eliminates the correlation between arbitrary pair of vectors in codebook, which is harmful for image extraction. These steps can dramatically improve the robustness of the watermark. However, the disadvantage of this method is that it doesn't resist to some kinds of geometric distortions, such as translation, flipping, rotation and scaling. Sanghyun Joo, Youngho Suh, Jaeho Shin, and Hisakazu Kikuchi in [2] propose a non-blind watermarking method for gray images that embeds a pseudo-random sequence (watermarks) into wavelet DC components. The DC area is not suitable for embedding because of severe visual degradation. This degradation problem is overcome by embedding watermarks into visually insensitive locations. Depending on the length of the watermarks, robustness and fidelity the wavelet decomposition level is determined. From nth level LL frequency sub band LL_n, reference LL frequency sub band LL_n' is obtained for watermark embedding. Chien-Pen Chuang, Cheng-Hung Liu, Yi-Tsai Liao, Huan-Wei Chi in [3] developed a robust digital watermarking technique to protect intellectual property right of digital image. A scale-invariant feature transform technique was employed for resisting geometric attacks and 2-D bar code was used for its high capacity and fault tolerance to enhance digital watermarking capacity. Besides, the Discrete Wavelet Transform (DWT) and discrete cosine transform (DCT) were joined to cope with noise problem and enhance perceptual transparency of watermarking image. The algorithm performance was presented for several attack models such as lossy compression, scaling, blurring and sharpening etc. The

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

experimental results proved the robustness of this mixed transformation technique on protecting intellectual property right of digital image. In [4] Yasuhiko Dote, Muhammad Shafique Shaikh have proposed a multiresolution based watermarking method in which coefficients of watermark are embedded into host image at different transformation level using secret key. For extraction, same key is used with inverse transformation at each level. Watermark is extracted by taking the mean value of the obtained watermarks. Method is tested for gray and color images and for various noises and proved to be robust.

In [5] Yingli Wang, Xue Bai and Shuang Van have presented a new digital image watermarking algorithm based on texture block and edge detection in the discrete wavelet domain. In the algorithm, the texture blocks are extracted after the edge detection for the original image with the canny operator by using the masking property of human visual system, in which the watermark is embedded adaptively both in the low-frequency sub-band and the high-frequency sub-band in the discrete wavelet domain. Texture blocks are selected for embedding the watermark as human visual system is not so sensitive to the noise in the strong texture based region. The technique shows good invisibility and robustness as well as the ability of resisting to geometric attacks. In [6], grayscale image watermarking embedding and detection is put forward by Qing Liu and Jun Ying, on the basis of digital watermarking principle and wavelet multiresolution analysis. It is an adaptive blind grayscale image watermarking algorithm. The embedded watermarking signal is processed by spread spectrum technology. The location of the embedded watermark and strength is automatically adjusted according to the characteristics of the original image, and watermark is adaptively added to the grayscale images. Experimental results show that the proposed algorithm enhances the anti- attack capability and the hidden nature of the image, improve the security of the watermarking detection, and has higher robustness to random noise attack, cutting and JPEG compression. In [7], a multimedia authentication and restoration scheme is proposed by Sidra Riaz and Sang-Woong Lee with the security of AES-128 ciphered watermarking and correlated watermarking. An encrypted or ciphered image embedding is done by modified version of Closest Point Transform (CPT) in a digital photograph. Image Authentication is done by locating the tempered areas and restoration is performed by correlated watermark on the tempered region of watermarked photograph. This correlated watermark is generated from the original image and is embedded in second and third level wavelet sub bands. The PSNR values are checked to evaluate the proposed watermarking technique. The results of PSNR, MSE, and SSIM show that the imperceptibility of proposed scheme is high compared to existing methods. In [8], Nagaraj V. Dharwadkar, B. B. Amberker & Avijeet Gorai proposed an effective, robust and imperceptible colour image watermarking scheme. This scheme embeds the watermark into cover image in (Red, Green, Blue) RGB space. The combination of Discrete Wavelet Transformation (DWT) and Singular Value Decomposition (SVD) of Blue channel is used to embed the watermark. The singular values of different sub band coefficients of Blue channel are modified using different scaling factors to embed the singular values of the watermark. The copy of the watermark is embedded into four sub band coefficients which is very difficult to remove or destroy. The combination of DWT and SVD increases the security, robustness and imperceptibility of the scheme.

In [9], Rahim Ansari, Mrutyunjaya M Devanalamath, K. Manikantan, S. Ramachandran, proposed a Digital Watermarking Algorithm using a unique combination of Discrete Wavelet Transform (DWT), Discrete Fourier Transform (DFT) and Singular Value Decomposition (SVD) for secured transmission of data through watermarking digital colour images. The singular values obtained from SVD of DWT and DFT transformed watermark is embedded onto the singular values obtained from SVD of DWT and DFT transformed colour image. Scaling and shift invariance property of DFT, rotation invariance property of SVD and robustness of DWT to compression are used to perform secure transmission of data through watermarking. In [10], to effectively resist geometric attacks and protect the security of information and copyrights of digital products, a digital image zero–watermarking algorithm based on geometric correction is proposed by Liu Peili, Tan Yuehui. The invariant centroid of an image is selected as a stable geometric reference point. The specific SIFT point which is the farthest to the reference point in image is obtained. Then geometric transformation parameters by changes of these two points' positions are calculated. To hide information in image, the image is divided into blocks and QR decomposition is applied to each block. After this the 2-Norm of the first row vector of matrix R in each block is extracted to form a sequence, and it is transformed into a matrix after binary quantization. Finally, the watermark and the matrix are XORed, and the result of the operation is the zero-watermarking. Experimental results show that the algorithm can effectively resist geometric attacks, with a high degree of correction accuracy and a strong robustness. By using the zero-watermarking method, the algorithm can also avoid the contradiction between invisibility and robustness. Sushila Kamble, Vikas Maheshkar, Suneeta Agarwal and Vinay Kumar Srivastava have proposed a wavelet based multiple watermarking of digital images in [11]. Watermark is embedded in LL sub band of host image. From LL sub band of host image, host vectors are created and two watermarks are embedded into these vectors by orthogonal projections. The use of multiple watermarks increases the embedding capacity of watermark as compared to the singular watermarking technique. Also, the spread transform disperse the watermark over a large number of frequencies providing robustness to volumetric as well as geometric distortions. Zhen Li, Kim-Hui Yap and Bai-Ying Lei have proposed a SVD-DCT based watermarking method in [12]. After applying SVD to the cover image blocks, DCT on the macro block comprised of the first singular values of each image block is taken. Watermark is embedded in the high-frequency band of the SVD-DCT block by imposing a

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

particular relationship between some pseudo-randomly selected pairs of the DCT coefficients.

Xi-Ping and Qing-Sheng Zhu [13] have proposed a wavelet based method using sub-blocks of image. Instead of applying wavelet transform on whole image, it was applied to local sub-blocks. These sub-blocks were randomly extracted from original image. Watermark was embedded into part of frequency coefficients of these sub-bands by computing their statistical characteristics. Krishnamoorthi and Sheba Kezia [14] proposed a watermarking technique based on orthogonal polynomial based transformation for copyright protection of digital images. A visual model was used to determine strength of watermarking. This visual model was used to generate Just Noticeable Difference (JND) by analysing low level image characteristics like texture, edges and luminance of cover image in polynomial based transformation domain. Ko-Ming-Chan and Long-Wen Chang [15] have proposed a watermarking system which embeds two different watermarks –robust and fragile into spatial and frequency domain separately. Robust watermark is embedded in wavelet coefficients of LL band whereas fragile watermark is embedded in least significant bits of watermarked image. Advanced encryption standard- Rijndeal block cipher was used to make watermarking technique public.

In [16],[17],[18],[19] we have proposed robust digital image watermarking algorithms based on various wavelet transforms generated from corresponding orthogonal transforms using Kekre's algorithm proposed by H. B. Kekre in [20]. These wavelet transforms are taken up to second level and further supported by DCT and SVD to increase robustness. However, this increases computational overhead.

In this paper, a new column DCT wavelet transform based compressed watermarking technique has been proposed. Full image transform is given by T*f*T' where T is transformation matrix and f is image to be transformed. Similarly inverse transform is obtained by T*F*T' where F is transformed image. Whereas, column transform of an image is given by T*f and inverse column transform is given by T'*F. Thus, taking column transform instead of full transform itself reduces the computation up to 50%. Performance of this technique is further explored under binary distributed and Gaussian distributed run length noise attack and image cropping attack. Section II explains DCT Wavelet Transform. Section III presents proposed technique and its performance for above mentioned attacks. Results are presented in Section IV. Section V concludes the paper.

II. DCT Wavelet Transform

DCT wavelet transform matrix is generated from orthogonal DCT transform using algorithm proposed in [1]. In the proposed method, host images of size 256x256 and watermark of size 128x128 have been used. Hence we require DCT wavelet matrix of size 256x256 and 128x128. These matrices can be generated using various combinations of component orthogonal DCT matrix. In the presented work, 256x256 size DCT wavelet matrix is generated using two 16x16 DCT matrices and 128x128 size DCT wavelet matrix is generated using 16x16 and 8x8 size DCT matrices. For an image f, transform of image is given by

$$\mathbf{F}=T*\mathbf{f}*T'$$

Where, T is orthonormal transform matrix. Inverse transform is given by

$$\mathbf{f}=T'*\mathbf{F}*T$$

The computational overhead of multiplication of matrices twice can be reduced to once by taking column transform or row transform of an image. Column transform of an image is given by

$$\mathbf{F}=T*\mathbf{f}$$

And inverse column transform is given by

$$\mathbf{f}=\mathbf{F}*T'$$

III. Proposed Technique

For the proposed technique, ten different host images (256*256) and five different watermarks (128*128) have been used as shown in Fig. 1 and Fig. 2 respectively.

The proposed technique can be divided into four phases: Watermark compression, watermark embedding, performing attacks on watermarked image and watermark extraction as shown in following Fig. 3.

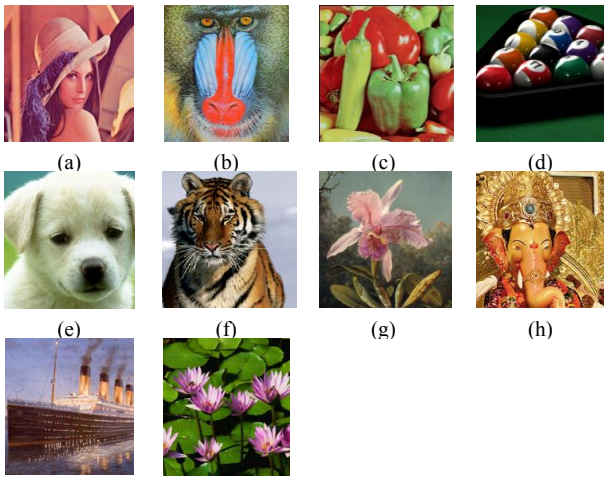


Fig. 1. Cover images used for experimental work (a) Lena (b) Mandrill (c)Peppers (d)Balls (e)Puppy (f)Tiger (g)Flower (h)Ganesh (i)Titanic (j)Waterlili

Fig. 2. Watermark images used for experimental work (a) Austral (b) Bear (c) CCD (d) Logo (e) NMIMS

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

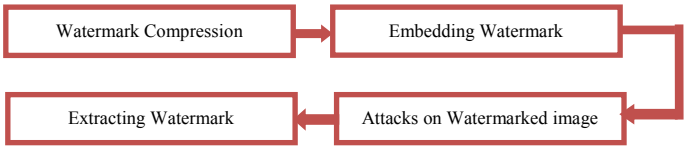


Fig. 3. Steps in proposed watermarking technique

A. Watermark compression:

In the proposed technique, watermark is not embedded as it is. Instead, it is first compressed using DCT Wavelet column transform [20]. Compression of watermark reduces the number of transformed coefficients to be embedded in host image. This reduces the number of altered transform coefficients in host image and hence the degradation caused in watermarked image. Steps for watermark compression are:

- Apply column DCT Wavelet transform on watermark image W.
- Eliminate 40, 48, 56, 64, 72 and 80 lowermost rows from transformed watermark to achieve compressed version of watermark with desired compression ratio.
- Reconstruct compressed watermark CW by taking inverse column DCT wavelet transform.

B. Embedding watermark:

- Apply column DCT Wavelet transform on host image.
- Normalize the compressed watermark CW to get normalized watermark CW_N.
- Replace the transform coefficients in the lowermost rows of host image by transform coefficients of normalized watermark.
- Take inverse column DCT Wavelet transform of host image to obtain watermarked image WI.

C. Attacks on watermarked image:

Three different types of attacks have been tested in the technique

- **Binary distributed Run length noise**

In this type of attack, noise with random run and discrete magnitude of -1 and 1 is generated with equal probabilities. This noise is then added to the watermarked image. From the attacked watermarked image, watermark is extracted by following the extraction procedure.

- **Gaussian distributed Run length noise**

In Gaussian distributed type of noise, once again noise of variable run with discrete magnitude in the range of -2 to 2 is generated and added to the image. Extraction procedure is followed to recover watermark from attacked watermarked image.

- **Cropping :**

Cropping of image is done at four corners with 16*16 size squares and 32*32 size squares.

D. Extraction of watermark:

- Take column DCT Wavelet transform of watermarked image.
- Extract the transform coefficients from lower rows of watermarked image.
- Denormalize these transform coefficients.
- Take inverse column DCT Wavelet transform of these denormalized transform coefficients to extract the watermark.

IV. Results

Fig. 4 on next page shows the result images for 'Lena' image and 'nmims' watermark for compression ratio 1.45 i.e. 40 lowermost rows are eliminated from transformed watermark image.

Fig. 4. Result images for 'Lena' host image and 'nmims' watermark. (a)host image (b) Original watermark (c) Compressed watermark (d) Watermarked image (e) Extracted watermark (f) watermarked image after binary distributed run length noise attack (g) Watermark extracted from watermarked image after binary distributed run length noise attack (h) Watermarked image after Gaussian distributed run length noise attack (i) Watermark extracted from watermarked image after Gaussian distributed run length noise attack (j) Watermarked image after cropping 16*16 squares at four corners (k) Watermark extracted from watermarked image subjected to 16*16 cropping attack (l) Watermarked image after cropping 32*32 squares at four corners (m) Watermark extracted from watermarked image subjected to 32*32 cropping attack.

Table I below shows readings for various MAE values for 'nmims' watermark embedded into ten host images. This watermark has been compressed using various compression ratios like 1.45, 1.6, 1.77, 2, 2.285 and 2.67. Watermarked images are further subjected to Gaussian distributed Run length noise attack.

Various MAE values shown in the table are MAE between watermark before compression and watermark after compression. Since ten host images are used for experimental work, average of MAE between host and watermarked image is also shown in table. Next MAE value shown in the table is between watermarked image and Gaussian distributed run length noise attacked watermarked image. The last column of table shows the average MAE between embedded compressed

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

watermark and watermark extracted from attacked watermarked image.

TABLE I. VARIOUS MAE VALUES FOR DIFFERENT COMPRESSION RATIOS USING 'NMIMS' AS WATERMARK AND GAUSSIAN NOISE DISTRIBUTED RUN LENGTH NOISE ATTACK ON WATERMARKED IMAGES

Watermark Image	Compression Ratio	MAE between original and compressed watermark	Avg. MAE between host and watermarked image	MAE between watermarked and attacked watermarked image	MAE between embedded compressed watermark and extracted watermark from attacked image
nmims	1.45	7.823	2.041	0.7463	68.03
	1.6	9.527	1.903	0.7463	65.73
	1.77	11.05	1.777	0.7463	63.78
	2	12.351	1.663	0.7463	63.02
	2.285	14.206	1.515	0.7463	58.83
	2.67	17.202	1.361	0.7463	56.38

From Table I it can be seen that as more compressed version of watermark is embedded, less is the MAE between host and watermarked image. This is due to less number of transform coefficients get altered in host image. Also MAE between embedded and extracted watermark gets reduced. However, MAE between watermarked and attacked watermarked image remains same irrespective of compression ratio.

Summary of average MAE between embedded and extracted watermark with different compression ratios for 'nmims' watermark embedded in ten host images against various attacks is given in Table II below.

TABLE II. AVERAGE MAE BETWEEN EMBEDDED NMIMS WATERMARK AND EXTRACTED NMIMS WATERMARK COMPRESSED WITH VARIOUS COMPRESSION RATIOS AGAINST BINARY DISTRIBUTED RUN LENGTH NOISE, GAUSSIAN DISTRIBUTED RUN LENGTH NOISE, CROPPING 16*16 AT FOUR CORNERS OF HOST IMAGE AND CROPPING 32*32 AT FOUR CORNERS OF HOST IMAGE ATTACKS ON WATERMARKED IMAGE

Compression Ratio	Type of attack			
	Binary run length noise	***Gaussian distributed run length noise***	***Cropping 16*16 at four corners of host image***	***Cropping 32*32 at four corners of host image***
1.45	Close to 0	68.03	1.29	4.275
1.6	Close to 0	65.73	1.42	6.118
1.77	Close to 0	63.78	2.18	7.421
2	Close to 0	63.02	17.37	36.33
2.285	Close to 0	58.83	1.15	3.94
2.67	Close to 0	56.38	1.29	5.26

From Table II, it can be said that, for binary distributed run length noise, the proposed technique outperforms. For Gaussian distributed run length noise, robustness increases as we go on embedding more compressed version of watermark. For 16*16 cropping attack, robustness is reduced as compression of watermark is increased till compression ratio 1.77. For compression ratio 2, robustness drastically falls to MAE value 17.37. However, for next higher compression ratio values, again better results are obtained but with increase in MAE between embedded and extracted watermark. For 32*32 cropping attack, similar trend is observed with increased error.

V. CONCLUSION

Compression of watermark before embedding it in the host image reduces the amount of information hidden in host image. This improves imperceptibility in watermarked image with minimal loss of information from watermark. Response of proposed technique is excellent for binary distributed run length noise with negligible MAE between embedded and extracted watermark. For Gaussian distributed run length noise, better performance is observed in terms of MAE as compression ratio is increased. However, for cropping (16x16 and 32x32 squares at four corners), performance drastically falls for compression ratio 2. Otherwise, for other compression ratios, better robustness is observed.

REFERENCES

[1] Hu Guan, Zhi Zeng, Shuwu Zhang, "A new DCT-based digital image watermarking algorithm", pp. 166-169.

[2] Sanghyun Joo, Youngho Suh, Jaeho Shin, and Hisakazu Kikuchi, "A new robust watermark embedding into wavelet DC components", ETRI Journal, Volume 24, Number 5, October 2002, pp. 401-404.

[3] Chien-Pen Chuang, Cheng-Hung Liu, Yi-Tsai Liao, Huan-Wei Chi, "A robust digital watermarking with mixed transform technique for digital image", Proceedings of the International Multiconference of Engineers and Computer Scientists, Vol II, March 2010, pp. 262-266.

[4] Yasuhiko Dote, Muhammad Shafique Shaikh, "A robust watermarking method for copyright protection of digital images using wavelet transformation", IEEJ Trans. Vol. 123 No. 2.

[5] Yingli Wang, Xue Bai, Shuang Van, "Digital image watermarking based on texture block and edge detection in the discrete wavelet domain", 2013 International Conference on Sensor Network Security Technology and Privacy Communication System (SNS & PCS), pp. 170-174.

[6] Qing Liu, Jun Ying, "Grayscale image digital watermarking technology based on wavelet analysis", 2012 IEEE Symposium on Electrical & Electronics Engineering (EEESYM), pp. 618-621.

[7] Sidra Riaz, Sang-Woong Lee, "Image authentication and restoration by multiple watermarking techniques with Advance Encryption Standard in digital photography", ICACT 2013, pp. 24-28.

[8] Nagaraj V. Dharwadkar, B. B. Amberker & Avijeet Gorai, "Non-blind Watermarking scheme for colour images in RGB space using DWT-SVD", 20 11 IEEE, pp. 489-493.

[9] Rahim Ansari, Mrutyunjaya M Devanalamath, K. Manikantan, S. Ramachandran, "Robust digital image watermarking algorithm in DWT-DFT-SVD domain for colour images", Proc. of IEEE International Conference on Communication, Information & Computing Technology (ICCICT), Oct. 19-20 2012, Mumbai, India, pp. 1-6.

[10] Liu Peili, Tan Yuehui, "Robust zero-watermarking algorithm based on invariant centroid", Proc. of IEEE International Conference on Computational and Information Sciences 2013, pp. 758-761.

[11] Sushila Kamble, Vikas Maheshkar, Suneeta Agarwal and Vinay Kumar Srivastava, "DWT-based Multiple Watermarking for Privacy and Security of Digital Images in E-commerce", Proc. of IEEE conference on International Conference on Multimedia, Signal Processing and Communication Technologies, 2011, pp. 224-227.

[12] Zhen Li, Kim-Hui Yap and Bai-Ying Lei, "A new blind robust image watermarking scheme in SVD-DCT composite Domain", 18th IEEE International Conference on Image Processing, 2011, pp. 2757-2760.

[13] Xi-Ping and Qing-Sheng Zhu, "A robust wavelet-domain watermarking algorithm for color image", Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, pp.13-16 August 2006.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[14] R. Krishnamoorthi, Sheba Kezia, "Image Adaptive Watermarking with Visual Model in Orthogonal Polynomials based Transformation Domain", IJICE, 5:2, 2009, pp. 146-153.

[15] Ko-Ming Chan, Long-wen Chang, "A Novel Public Watermarking System based on Advanced Encryption System", IEEE Proc. of 18th International Conference on Advanced Information Networking and Application, 2004.

[16] H. B. Kekre, Tanuja Sarode, Shachi Natu, "Performance Comparison of DCT and Walsh Transforms for Watermarking using DWT-SVD", International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013, pp. 131-141.

[17] H. B. Kekre, Tanuja Sarode, Shachi Natu, "Hybrid Watermarking of Colour Images using DCT-Wavelet, DCT and SVD", International Journal of Advances in Engineering and Technology, vol. 6, Issue2, May 2013, pp. 769-779.

[18] H. B. Kekre, Tanuja Sarode, Shachi Natu, "Robust watermarking using Walsh wavelets and SVD", International Journal of Advances in Science and Technology, Vol. 6, No. 4, May 2013, pp. 8-23.

[19] H. B. Kekre, Tanuja Sarode, Shachi Natu, "Performance Comparison of Wavelets Generated from Four Different Orthogonal Transforms for Watermarking With Various Attacks", International Journal of Computer and Technology, Vol. 9, No. 3, July 2013, pp. 1139-1152.

[20] H. B. Kekre, Archana Athawale, Dipali Sadavarti, "Algorithm to generate wavelet transform from an orthogonal transform", International Journal of Image Processing, Vol.4, Issue 4, 2010, pp. 444-455.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

CONTENT BASED IMAGE RETRIEVAL THROUGH DISTANCE SIMILARITY METRICS

Praveen Mishra
Department of Computer Science
VNS Institute of Technology Bhopal, India
mishrapraveen63@gmail.com

Prof. G.S Prajapati
Department of Computer Science
VNS Institute of Technology
gsprajapati1234@gmail.com

Abstract:

Searching Test Image from Image databases using features extraction from the content is currently an active research area. In this work we presents novel feature extraction approaches for content-based image retrieval when the query image is color image. To facilitate robust man-machine interfaces, we accept query images with color attributes. Special attention is given to the similarity measure with different distance matrices properties since the Test Image and Object Image from database finding the distance measuring. Several applicable techniques within the literature are studied for these conditions.

The goal is to present the user with a subset of images that are more similar to the Object Image. One of the most important aspects of the proposed methods is that the accuracy measurement of the different database images. This significantly improves the feature extraction process and enables the methods to be used for other computer vision applications.

Keyword: Color image, Image retrieval, Euclidean metrics, Manhattan metrics and Correlation.

1. INTRODUCTION:

Content Based Image Retrieval (CBIR) is a set of techniques for retrieving semantically relevant images from an image database based on automatically derived image features [1]. The computer must be able to retrieve images from a database without any human assumption on specific domain (such as texture vs. non texture or indoor vs. outdoor).

One of the main tasks for (CBIR) systems is similarity comparison, extracting feature signatures of every image based on its pixel values and defining rules for comparing images. These features become the image representation for measuring similarity with other images in the database. Calculating the difference of the feature components for compared images.

Early CBIR methods used global feature extraction to obtain the image descriptors. For example, QBIC [2], developed at the IBM Almaden Research Center, extracts several features from each image, namely color, texture and shape features. These descriptors are obtained globally by extracting information on the means of color histograms for color features; global texture information on coarseness, contrast, and direction; and shape features about the curvature, moments invariants, circularity, and eccentricity. Similarly, the Photo-book-system [10], Visual-Seek [11], and Virage [12], use global features to represent image semantics.
Image database is one of the centerpieces for distributed multimedia systems. In general, an image database system is an intelligently combination of the following three major components:

- image processing
- storage, retrieval, and management
- user interface

The first component deals with processing for the extraction of information from original images while the second component is providing efficient tools for storing, retrieval and managing of image data. Querying the database needs a simple and user friendly interface which is the responsibility of the third component.

2. RELATED WORK:

2.1 PIXEL DOMAIN TECHNIQUES:

Typically, visual indexing techniques are based on features which are extracted directly from the pixel domain. These include visual features that the human visual system can easily recognize. Working in this domain is more intuitive and does not require us to transform and re-transform the images at the store and retrieving stages,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

respectively. Here, we review the most important image clues in the pixel domain, which are color, texture, shape, spatial relation, and sketch.

2.2 COMPRESSED DOMAIN TECHNIQUES:

The large volume of visual data necessitates the use of compression methods. Multimedia databases usually store the visual content in compressed form and most images are stored using existing compression techniques. In order to reduce the cost of decompression of the image data and applying pixel domain techniques, it may be more efficient (if applicable) to index visual information in the compressed form. This approach often has a lower cost for computing and storing the indexes.

Compressed domain indexing is broadly classified into two main categories: spatial domain, and transform domain techniques. The major techniques in the first category are vector quantization [3] (VQ) and fractals. The second category is generally based on image transformation techniques including DFT [4] (discrete Fourier transform), DCT (discrete cosine transform), sub bands/wavelets transforms, Haar Transforms, and KLT (Karhunen-Loeve transform).

3. PROPOSED WORK:

The goal is to present the user with a subset of images that are more similar to the Query image. New affine transform invariant feature extraction techniques are proposed to improve retrieval performance and reduce the extraction and search times. The techniques are tested both generally for multi-component images and particularly for any pictures. The solutions are discussed for each specific application. Finally, content-based image retrieval [5], which explores image retrieval from databases using different distance metrics, is investigated on an individual basis.

Two different approaches based on CBIR i.e. Pixel domain based and another one is compressed domain based for general image retrieval. Here, the database images consist of multiple complex images within an inhomogeneous background. One of the methods is an improved version of another, which increases retrieval performance. Both techniques exhibit their better task perform better result.

Our approach is to retrieve the image from the frequency domain method. By this method we firstly apply the color image conversion i.e. in the form of RGB.We must have to apply the color conversion technique first with the help of color model i.e. HSV (Hue, Saturation and value).The MATLAB code is so much helpful for us for converting the image color in HSV format. After converting we are selecting the blocks the picture which is helpful for finding the distance measure. Here we apply the similarity measure between two images with the help of different metrics such as Euclidean metrics so on.

Firstly we take two images and make the blocks into it which requires to measure the distance of metrics by image 1 find the Euclidean distance that is called Eudclean1. Now again find the distance of another image we get the Euclidean distance that is Eucledean2. After it we can find the distance with the help of these two distances i.e. called Manhattan metrics.

4. ALGORITHM:

4.1. Preprocessing
Step1:- Input various object images
Step2:- Create 4*4 Block matrix
Step3:- Calculate Mean of Block Matrices
Step4:- Concatenate all Block Matrices
Obtain from step 3

4.2. Feature Extraction
Step5:- Convert Block Matrices “f and g” RGB from space to HSV from space.
% (where f and g are block matrices of query and object image)
Step6:- Extract feature vector from HSV space

4.3. Similarity Measure
Step7:- Calculate Euclidean Distance then get Euclidean 1 (D_2)

$$D_2(f, g) = l_2(f, g) = \sqrt{\sum_{i=1}^{N} (fi - gi)^2}$$

Step8:- While (T_i= =O_i) %(where T_i is the test image and O_i is the object images)
{
Repeat all procedure on another test image and object images

Step9:- Now we have ‘N’ object image and its Euclidean distance metrics.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Step10:- Then Calculate Manhattan Distance (D_1)

$$D_1(f, g) = l_1(f, g) = \sum_{i=1}^{N} |f_i - g_i|$$

}

Step11:- Evaluate Accuracy as
Accuracy= (Tp+Tn) / N

N=Tp+Fp+Tn+Fn
% whre Tp = Count of true positive
Fp = Count of false positive
Tn = Count of true negative
Fn = Count of false negative.
End

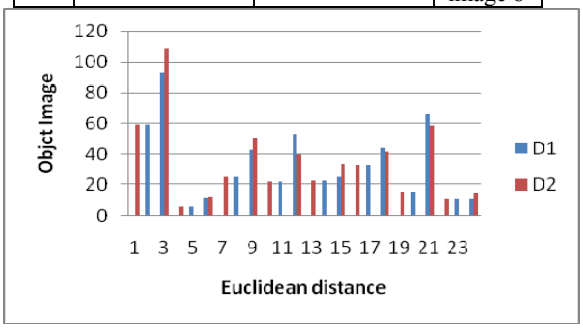
5. DATA SET AND RESULTS:

TEST IMAGE DATA:

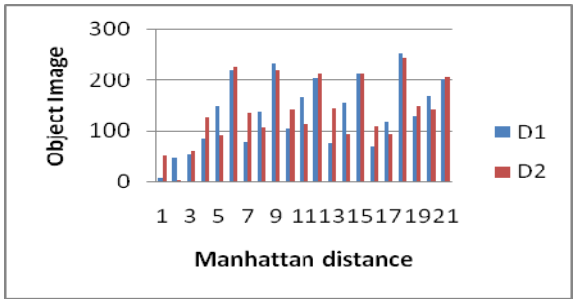
			
Running Bicycle(T_1)	Desert(T_2)	Daughter (T_3)	
 Object image 1	 Object image2	 Object image3	 Object image4
 Object image 5	 Object image 6	 Object image7	 Object image 8
 Object image 9	 Object image 10	 Object image 11	 Object image 12
 Object image 13	 Object image 14	 Object image 15	 Object image 16
 Object image 17	 Object image 18	 Object image 19	 Object image 20

Image Categorization through Evaluation of Object images and Testimage1

Object Images	Testimage1	Testimage2	Testimage3
	Object image 1		
		Object image3	
			Object image 6
	Object image 7		
		Object image 9	Object image 4
	Object image 2	Object image 5	Object image 8

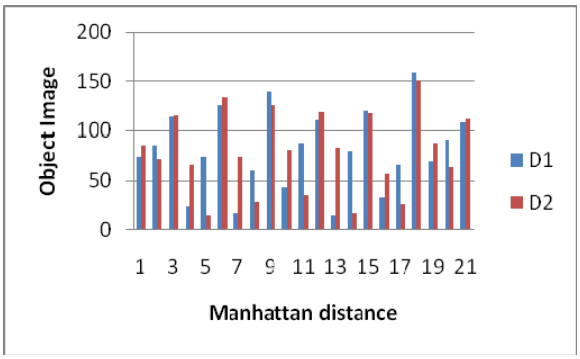


Graph show the Euclidean distance D1 and Object Image D2



Graph show the Manhattan distance and Object Image distance which show the similarity measure between TestImage2 and Object Images

http://sites.google.com/site/ijcsis/
ISSN 1947-5500



Graph show the Manhattan distance TestImage2 and different Object Image distance which show the similarity measure between TestImage2 and Object Images

6. CONCLUSION:

In conclusion, this thesis has presented several novel techniques for invariance feature extraction used in the CBIR. They achieve significant improvement in retrieval accuracy. The proposed techniques describe the similarity match between Object Images and Test Images. The Test images of the color feature must match with the color feature of the Object Images. The proposed methods were based on the distance similarity matrices of images. We use different matrix for such as Euclidean, Manhattan, chebyschib this discipline has been showed to be suitable for general images.

7. FUTURE WORK:

Possible improvements and further studies on the proposed methods are addressed below

- In the distance similarity measure method, the number of same images could be assigned dynamically. This means, we are calculating the distance with the pixels for Euclidean1. This would improve the performance but might slow down the feature extraction process. Moreover, variation in feature vector dimension makes the image matching complex. One solution to this problem can be partitioning the database images according to the respective feature vector dimension. In the matching process, the features extracted from the query image need to be compared only with the similar dimension features of the database images.

- The Manhattan distance must be calculated in above then we correlate it, which is described above, can be applied to the Eigen vector method as well. Here it indicates dynamically assigning an appropriate sector size for different images based on the existing density of pixels .The same problem of complex matching will arise, which could be solved as described.

REFERENCES:

[1] Szántó, B. ; Pozsegovics, P. ; Vámossy, Z. ; Sergyán, S. "Content-based image retrieval system using sketches" Applied Machine Intelligence and Informatics (SAMI), 2011 IEEE 9th International Symposium .

[2] David gorisse, matthieu cord, and frederic precioso "locality sensitive hashing for chi2 distance" IEEE transactions on pattern analysis and machine intelligence, vol. 201, no. 56, june 2010.

[3] PinderjeetKaur , KiranJyoti"Implementation of CBIR System for CAD Jewellery Images Using PCA" International Journal of Scientific & Engineering Research Volume 4, Issue 2, February-2013 ISSN 2229-5518

[4] K.P. Ajitha Gladis, K.Ramar, "A Novel Method for Content Based Image Retrieval Using the Approximation of Statistical Features, Morphological Features and BPN Network", IEEE computer society ICCIMA 2007 ,Vol. 148 , PP. 179-184

[5] Chunbo Zhu]*, Xiaoyan Sun2, Feng Wu2, and Houqiang Li, "Video Coding With Spatio-Temporal Texture Synthesis", IEEE ICME 2007, Vol 07, PP. 113-115.

[6] Tienwei Tsai, Te-Wei Chiang, and Yo-Ping Huan, "Image Retrieval Approach Using Distance, Threshold Pruning", IEEE Trans. On Image Processing 2007, Vol.12,PP.241-249.

[7] C. Wang, X. Sun, F. Wu, and H. Xiong, "Image compression with structure-aware

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

in painting," in Proc. ISCS 2006, Vol 19,pp. 1816-1819.

[8] Hakinii HACID, Abdelkader Djalmiel ZIGHEI), "Content-Based Image Retrieval Using Topological Models" IEEE Trans. On image processing 2006 ,Vol 28, PP.308-311

[9] Yong Xia, (David) Dagan Feng, and Rongchun Zhao, "Morphology-Based Multifractal Estimationfor Texture Segmentation", IEEE ICCIMA 2006, Vol. 15, PP .1125-1130

[10] S. Newsam and C. Kamath, "Comparing shape and texture features for pattern recognition in simulation data," in SPIE Electronic Imaging, San Jose,CA, January 2005, pp. 106–117.

[11] MPEG-7 Home Page http://www.darmstadt.gmd.de/mobile/MPEG-7/index.html

[12] Anil K. Jain, "Fundamentals of Digital Image Processing", Pearson Education PHI Publishers, 2004

[13] Pentland, A., Picard, R. and Sclaroff S., "Photobook: Contentbased manipulation of image databases", International Journal of Computer Vision 1996, Vol 18(3), pp.233–254

[14] Smith, J.R., and Chang, S.F., "Single color extraction and image query," In Proceeding IEEE International Conference on Image Processing, 1997,Vol 29, pp. 528–531

[15] Gupta, A., and Jain, R., "Visual information retrieval," Comm. Assoc. Comp. Mach, 1997., Vol 40(5), pp. 70–79

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

EVOLUTIONARY NEURAL NETWORK MODEL FOR DYNAMIC CHANNEL ALLOCATION IN MOBILE COMMUNICATION NETWORK

Ojesanmi Olusegun and Ugege Peter

Department of Computer Science
Federal University of Agriculture
Abeokuta, Nigeria.

Abstract
Mobile communication aims at transmitting multimedia data and at the same time provides guaranteed quality of service (QoS) to all the applications. The challenge is to develop an efficient allocation scheme for assigning resources without compromising the QoS. In meeting this challenge, this paper proposes an evolutionary neural network approach with dynamic allocation to utilize frequency spectrum efficiently and to reduce call blocking probabilities.

Keywords: Neural Network, Mobile network, call blocking probabilities, dynamic channel, QoS, Multimedia.

1.0 INTRODUCTION

The rapid growth in recent demand for mobile telephone services and the limited nature of the electromagnetic spectrum allocated for this purpose makes solving the problem of channel assignment a more critical one. The channel assignment problem (CAP) is to efficiently allocate channels or frequencies to each cell in the cellular network, while satisfying the electromagnetic interference constraints. The new generations of wireless access system offers wideband services, such as multimedia thus moving the trend towards a larger number of cellular users, the reality here is that the available frequency spectrum for allocation to these users becomes more limited. Consequently, the effective use of channel frequencies becomes more and more important. The CAP is classified as an NP-complete problem [63, 76, 41, 60, 59] meaning that as the size of the problem increases, the time required to solve the problem increases in an exponential form and not in a polynomial form. That is it requires a search space which exponentially increases in order to identify the best channel combination among the cells as the numbers of the base stations and users increases. Most approaches to the CAP have been developed under the assumption that the base-stations are regularly distributed over the service area. As a result, the channel assignment at each cell should consider the Co Channel Interference effect only on these adjacent cells. However, a spectral efficiency per unit area obtained from macro-cellular systems is not sufficient to support the exponentially increasing demand for wireless data service because the area covered by one macro-cell is too large [21].

Due to the radio interference between frequency spectra, [70] stated that the channel assignment must fulfill three constraints which are normally determined by the characteristics of the radio frequency propagation and spatial density of the expected traffic requirements. The constraints are: (i) The Co-Channel Constraint (CCC): The same channel cannot be assigned to a pair of cells within a specified distance simultaneously. (ii) The Adjacent Channel Constraint (ACC): Adjacent channels cannot be assigned to adjacent cells simultaneously. (iii) The Co-Site Constraint (CSC): The distance between any pair of channels used in the same cell must be larger than a specified distance.

Channel allocation is viewed as a large-scale dynamic optimization problem with multiple goals and constraints in a stochastic environment. Reinforcement learning approach [47], evolutionary algorithms [36, 29], heuristic oriented search [28], neural network and other computational intelligence approaches [9, 51, 18, 70, 16, 10, 11, Kendall and Mohammad, 2004,36, 28] have been employed alone or in combination with others to solve the CAP. The advantage of these approaches is that they consider environmental dynamics and the intelligence strategies can be adapted in accordance with information obtained in the behavior of the environment. Among the evolutionary algorithms, genetic algorithms (GA) play an important role with their exhibition of implicit parallelism associated with the ability to effectively explore information over large search spaces. GAs have becomes very efficient in solving any difficult optimization task and NP-complete problems. In this work we are

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

combining the search ability of Genetic Algorithm with Neural network model to address the CAP.

1.1 RESEARCH MOTIVATION

The effect of congestions in mobile communication networks is quite obvious and reflects the inability of such network to provide satisfactory quality of service (QoS) to the network subscribers. Inefficient allocation of the limited channel resources to cells in the networks remains a challenge and a major cause of congestions in mobile communication networks. The need for a more efficient scheme to address the channel allocation problem in mobile communication network drives the motivation for this work.

1.2 PROBLEM STATEMENT

Channel allocation is viewed as a large-scale dynamic optimization problem which is NP-complete in nature. The wideband kind of services such as multimedia being offered by new generations wireless access system makes the CAP more critical. Existing allocation schemes have not been able to address this problem effectively due to its NP-complete nature. Computational intelligence approaches have however proved effective for this type of problems and have been applied alone or in combination with others to address such problem. In this work we are combining the advantages of two computational intelligence techniques: Genetic algorithms with its ability to quickly produce a near optimal solution over a large search space and neural network with the ability to easily produce a local optimal solution within a small search space to solve the CAP with the aim of arriving at an efficient optimal solution to the problem.

1.3 RESEARCH OBJECTIVES

The specific Objectives are to:

i. Survey the different approaches to the CAP in mobile communication networks
ii. Develop an efficient model for optimal channel utilization in cellular mobile network
iii. Simulate the model in (ii) above
iv. Implement and assess the proposed model.

1.4 RESEARCH METHODOLOGY

i. State of the art review of existing channel allocation models
ii. Development of an optimal channel allocation model using evolutionary neural network approach
iii. Simulation of the model in (ii) using MATLAB
iv. Performance evaluation of the proposed model

2.0 LITERATURE REVIEW

The challenge of congestion control in wireless mobile networks is not novel. Both intelligent based and non-intelligent based approaches have been adopted to address the challenges. Various channel assignment schemes have been studied widely to find better ways to assign channels to calls and to achieve higher level of channel reuse. The fixed channel allocations (FCA) [4, 19, 73, 46] allocate channels permanently to each cell based on predetermined estimated traffic. FCA scheme is simple but it does not adapt to the changes in traffic conditions. In [38, 74], there is no permanent allocation of channels to cells. Rather, the whole set of available channels is accessible to all the cells, and the channels are assigned on a call-by-call basis, in a dynamic manner. Dynamic methods have better performance than FCA methods for light to medium traffic load [77, 74, 33]. Most of the proposed DCA algorithms are based on heuristics, and do not guarantee an optimal solution. In addition, many existing DCA schemes consider a simplified problem with only co-channel constraints [19, 38, 34, 13, 40, 65]. Hybrid allocation [15, 76, 78, 2, 61] combined the features of both fixed and dynamic techniques to overcome the drawbacks. Hybrid allocation performs better than other channel allocation schemes in terms of call dropping and blocking by using hot-spot notification and central pool [37, 71, 74].

In [9] an effective and efficient Hybrid Discrete Binary Differential Evolution Algorithm (HDB-DE) was presented for the solution of Dynamic Channel Assignment Problem. The features of HDB-DE help in achieving fast convergence and good solution. Their approach took care of the soft constraints as well as some hard constraints and hence focuses search only in the feasible regions of the search space resulting in fast convergence. The obtained results and their comparison with those obtained by the other methods established the superiority of HDB-DE over other methods and indicate its suitability for the solution of even bigger instances of dynamic channel assignment in cellular networks resulting in lower call rejection probability, higher capacity utilization and a good quality of service. Some of the hard constraints were not however considered. [57] proposed a GA-based channel allocation scheme that fairly supports best-effort users in a microcellular system was proposed. He had argued that the conventional scheme does not properly support the users far from the serving base-station, because it chooses the channels which improve the performance of the near users. To solve this problem, the proposed scheme finds the channel set maximizing the far users' performances while allowing a pre specified loss of overall performance. They did not consider any of the soft constraints. In [7], an optimization algorithm based on GA is proposed to solve the channel assignment problem in the cellular mobile network to achieve lower call dropping or call blocking probability and capable of mimicking the evolutionary process in nature in order to optimize the channel assignment problem was proposed. Its characteristics to evolve through generations and to select the fittest optimum chromosomes enable it to be self-optimized from

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

generation to generation. The proposed GA-based algorithm was able to perform the channel optimization smoothly with minimum level of calls blocked. Only hard constraints were considered in this work. [29] developed a reliability-based model that uses the GA to optimize the reliability in mobile computing network was proposed. The proposed model was an effective approach to make the network connections more reliable. It has been observed that the well-managed and efficient usage of the better channels (with lower failure rates) and delivering them to the mobile hosts greatly increases network reliability. They claim this to be the first work that conducted the CAP based on reliability values. None of the hard or soft interference constraints were considered.

In [53, 54, 55, 56], a more effective channel allocation scheme is introduced. The Multiple Cell-Multiple channel Multimedia Unbalanced DCA (MMDCA), that uses different cells in the neighbor. Thus, multiple channels can be assigned from different cells for supporting a video call. With this technique, low call dropping and blocking probability can be achieved due to multiple channel availability among different neighbor cells. They only considered some of the hard constraints. [82] proposed two efficient integer linear programming (ILP) formulations, for *optimally* allocating a channel (from a pool of available channels) to an incoming call such that both "hard" and "soft" constraints are satisfied. Their first formulation, ILP1, does not allow channel reassignment of the existing calls, while their second formulation, ILP2, allows such reassignment. Both formulations can handle hard constraints, which includes *co-site* and *adjacent channel* constraints, in addition to the standard *co-channel* constraints. The simplified problem (with only co-channel constraints) can be treated as a special case of their formulation. In addition to the hard constraints, they also considered soft constraints, such as, the *packing condition, resonance condition,* and *limiting rearrangements*, to further improve the network performance. [71] proposed a scheme carried in the general context of channel allocation issue involved in cellular networks at heavy load situations. They illustrated the efficiency and flexibility of hybrid channel allocation algorithm along with a slight modification in conventional algorithm by introducing the channel retrieval phase which retrieves back the channels with BS, borrowed from MSC and are unused and allocates the same to some BS requiring it and thus resolving the limitation of MSC of being exhausted of communication bandwidth to serve further calls in heavy load. This work gave better result than the simple hybrid channel allocation scheme that was proposed in [37] in terms of call dropping probability. Both schemes did not consider the interference issues. In [62], a new Genetic Algorithmic approach to solve Static Channel Allocation Problem was proposed. Their main goal was to assign channels to cells with efficient usage of bandwidth. The methodology consists of new crossover and mutation techniques, based on the reuse distance of channels. In crossover the selected channels from an individual will be used for generating a new child for next generation through which utilization and reassign of channels will be efficient. In mutation selecting the channel, checks within individual and reassigns the channel based on the reuse distance. Soft interference constraints were not considered in this work. [81] proposed a novel integer linear program (ILP) formulation that jointly optimizes channel allocation and power control for incoming calls, based on the carrier-to-interference ratio (CIR) was presented. In the approach a hybrid channel assignment scheme is used, where an incoming call is admitted only if a suitable channel is found such that the CIR of all ongoing calls on that channel, as well as that of the new call, will be above a specified value. Their formulation also guaranteed that the overall power requirement for the selected channel will be minimized as much as possible and that no ongoing calls will be dropped as a result of admitting the new call. They did not consider any of the soft constraints. In [68] a channel allocation model is proposed based on mutual exclusion algorithm for distributed system. Requests time stamped with Lamport's clock are sent by a MSS to its neighboring MSSs to transfer a free channel to support a communication session. No consideration was given to the interference constraints. [29] proposed a GA based fault tolerant channel allocation to minimize the average number of blocked hosts and handoff failures in mobile computing environment. They consider handoff problem by using channel reservation and channel borrowing techniques. This technique was not based on any consideration of the interference constraints. [45, 30] have applied modified GA for channel allocation for DCA. A new genetic operator was introduced for improving a simple GA. No consideration was given to the soft constraints.

In [30] a model was developed that utilizes the Global Positioning System (GPS) data for tracing the hosts' likely movements within and across the cells and allocates the channels to the mobile devices accordingly. The allocation of the channels to the mobile hosts is deterministic in the sense that the decision of the channel allocation is based on the realistic data received from the GPS about the hosts' movements. The interference constraints were not considered. [18] proposed procedures to generate lower bounds on the number of total frequencies while respecting all the channel constraints. In [70], a sequential heuristic algorithm based on the first algorithm introduced by [18] was developed and applied to several problems, where the values of total frequencies in solutions are shown without any actual assignment results. Unfortunately, none of the existing heuristic algorithms can find the lower bound solutions in all of the 13 instances for [70] benchmark problems. Lemmas formulated by [18], have been extensively used to test algorithms

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

proposed to solve the Channel Allocation Problem (CAP) [16, 79, 14]. Among basic strategies to solve the CAP include the Frequency Exhaustive Assignment (FEA) and Requirement Exhaustive Assignment (REA) heuristics. Considering a recently ordered set of the available frequencies, the FEA strategy assigns calls to the minimum available frequencies, while respecting the constraints. On the other hand, using REA, the lowest available frequency is assigned to the maximum number of calls that is possible respecting the constraints. Many works have used combinations of FEA or REA together with other techniques, in order to find near optimal solutions. In [70], three local search algorithms to solve the CAP, named CAP1, CAP2 and CAP3 were developed. The best result was obtained by CAP3. None of the soft constraints was considered. [51] proposed a two-stage algorithm for channel assignment problem by combining sequential heuristic algorithms into a parallel neural network algorithm. In 1991, [35] used the continuous Hopfield network, where the output of each neuron *Vi* was a fixed function *f* of the internal state *ui*, i.e., $Vi=f(ui)$, where $f(x)=1/2(1+\tanh(\lambda x))$. The neural-network model of [35] required a large number of iterations in order to reach the final solution, and there were also difficulties in finding the proper values for and the parameters in the interconnection weights and energy function. He considered only co-channel and co-site interference in his neural-network model.

In [49], a centralized dynamic allocation was employed with ticket scheduling technique to handle calls of the same priority. Buffers at various locations in the network, is introduced to handle the blocked real-time and non-real-time handoff in order to keep the quality of service of handoff calls. Before allocating channels to calls, if call request for channel arrive at the same time, the scheme gives priority to hot spot calls before cold spot calls if the cold spot call is non-real-time. However, if call comes from the same cell and of the same type, ticket scheduling is also invoked to serve the calls. This is an important step forward from other conventional static strategies that may be employed by network operators. New non-real-time traffic is assumed to be delay-tolerant and the capability of buffering the delay-insensitive real-time and non-real-time traffic is added. The scheme prevents traffic congestion because of its implementation at the base station subsystem of the network. The scheme is modeled with a multidimensional Markov chain. None of the interference constraints was considered. In [50] a dynamic multiple-threshold channel allocation scheme for cellular wireless networks with one-level buffer was proposed. The thresholds were dynamically adjusted according to the current network traffic situation and QoS status. The single buffer was used to queue both real-time and non-real-time handoff calls in case of unavailability of channel at the arrival of such calls. Their objective was to provide QoS by keeping the real time and non-real-time handoff connection dropping probability below the predefined bound even under a network congestion situation. Also to maintains the relative priorities among real-time calls and non-real-time calls in terms of the new connection blocking probability according to their traffic profiles and instantaneous traffic situations. The work was not based on any consideration of the interference constraints.

[1] implemented a genetic algorithm approach to address the frequency allocation problem in a wireless network. They did not consider the interference constraints in their implementation. In [10] an evolutionary approach for frequency assignment problem in cellular radio networks was presented. [76] presented an evolutionary strategy was used to solve the CAP for a hybrid channel allocation, with consideration given to a smaller set of constraints. The co-site interference constraint was left out of consideration. [16] proposed a three-stage genetic algorithm, combined with FEA strategy to solve the CAP. In [11], a simulated annealing meta-heuristic was applied to solve the CAP, but the compatibility matrix, used in that work, treated all interference relations between base stations with equal strength. A different set of instances and a graph coloring approach were used. Same approach was also adopted in [12, 43, 20]. None of the soft constraints was considered

In (Fernando and Fapojuwo, 2002) a Viterbi-like algorithm (VLA) was proposed to solve the CAP. The VLA is based on the original Viterbi algorithm used for information decoding in digital communication systems. The VLA uses as metric, instead of the Hamming distance, the excess frequency factor which is defined by the adjacent and co-channel constraints, compared with the difference of the frequencies assigned to a pair of cells. Soft constraints were left out of consideration in this work. To find the solution to the CAP (Kendall and Mohammad, 2004) described a hyper-heuristic strategy. The hyper-heuristic is a problem independent strategy (high level), that manages a set of low level heuristics (LLH), defining the borderline of the data flow. It combines a greedy constructive heuristic to generate the initial solution, and then a group of LLH is used to improve the solution quality. An algorithm chooses which heuristic will be used each time. The work did not consider the soft constraints. In (Liu et al., 2000) and (Gomes et al., 2001) GRASP (Greedy Randomized Adaptive Search Procedure) approach was used to solve the CAP, both using a graph coloring model. In (Liu et al., 2000) GRASP was combined with simulated annealing in the local search phase. In (Gomes et al., 2001) GRASP was combined with Path Relinking as a strategy for local search. No consideration of the interference constraints was considered.

2.1 Genetic Algorithm

Genetic Algorithm is one form of evolutionary algorithm (EA) which originates from

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

the principal of natural selection and survival of the fittest, and constitutes an alternative method for finding solutions to highly-nonlinear problems, by exploring multimodal solution space (Chia et al., 2011). It is a search algorithm based on the mechanics of natural selection, genetics and evolution. Figure 1 shows the genetic flowchart. They have been used in a variety of applications. In (Sancho et al., 2008) a hybrid genetic algorithm was used for optimal switch location in mobile communication networks. (Chandralekha and Praffula, 2010) used genetic algorithm to minimize the number of handoffs in heterogeneous wireless networks. (Aizaz et al., 2012) implemented genetic algorithm in dynamic channel allocation. Genetic algorithms work with a large *population* of solutions with each solution represented as a *chromosome*. A population is a collection of chromosomes (solutions). As in genetics, a chromosome contains several *genes*. Chromosomes are often represented in binary numbers with each binary bit corresponding to a gene. From the population of solutions, GA selects the best possible solution on the basis of a threshold or *fitness function* which is unique for every optimization problem. The fitness of each chromosome in the population is measured, and the best chromosome is selected. GA ensures a quicker convergence to the near-optimal solution (Rincy and Jude, 2011).Formulating the fitness function is the most crucial part of the genetic algorithm development. The fitness function is designed such that the best chromosome corresponds to the one with the least fitness value. For simplification the chromosomes and their fitness values can be represented in a matrix A as shown below.

$$A = \begin{pmatrix} String\ 1 & & f(x1) \\ \vdots & & \\ String\ n & & f(xn) \end{pmatrix}$$

Where string 1 …string n represent the chromosome and f(x1) …f(xn) represent the fitness value. Once the fitness values have been generated for each chromosome, half of the population size with the best fitness values (least fitness values) is selected and used to generate new chromosomes through the process of crossover and mutation. Crossover is the process where two chromosomes known as the parent chromosomes are combined to form two new chromosomes known as the child chromosomes. Crossover can be either one-point or two-point crossover. In Mutation only one parent is involved in forming a new chromosome. Some random genes are selected for mutation or change.

2.2 Neural network

Neural networks are algorithms for optimization and learning based loosely on concepts inspired by research into the nature of the brain. They generally contain the following components as shown in figure 2:

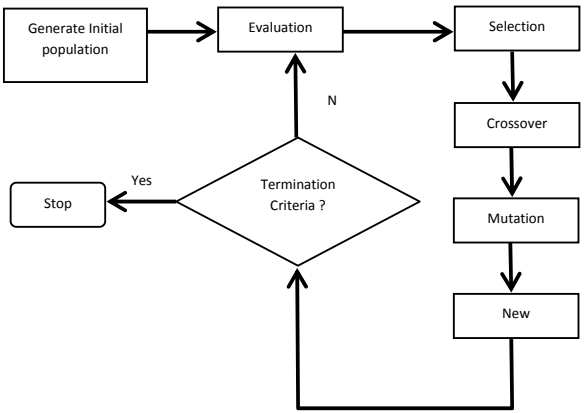


Figure 1: General flowchart of Genetic Algorithm

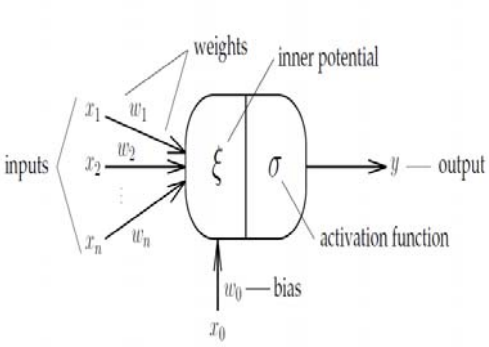


Figure 2: A neural network Model (David, 2010)

- inputs – Neuron consists of n inputs representing dendrites of the biological model. We can formally denote them as a vector (x1; x2; : : : ; xn).
- weights – Each input is weeded with its synaptic weight. The weight simulates the permeability of the membrane. The bigger the weight is, the more permeable membrane would be in the corresponding biological neuron. Therefore we can write the weights as a vector of n numbers (w1,w2,…,wn).
- bias – According to biology, the neuron provides an output when the threshold is reached. Negative value of the threshold t is represented as a weight of a special input in the artificial model called bias. It means $w_0 = -t$. For its formal input x_0 holds that at any time $x_0 = 1$. Hence the value of bias w_0 is fully used when computing an inner potential.
- inner potential – The weighted sum of all inputs (including the bias) is called inner potential. Formally: it is given as

$$\sum_{i=0}^{n} w_i x_i$$

- activation function – Inner potential is evaluated by an activation function.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- output – The value of the activation function is denoted by y. It is the output of the neuron

2.3 PROBLEM FORMULATION AND REPRESENTATION

The channel assignment problem being addressed in this work is based on a common model in which the service area of the system is divided into a large number of hexagonal cells and every user located in one cell. Our channel allocation scheme under consideration is the Dynamic Channel Allocation scheme. When a user service requests comes to the system, a channel or frequency spectrum is assigned there to provide the communication service. This channel assignment must satisfy certain constraints to avoid radio interference between channels, avoid congestion in the network and provide a good quality of service to users (see figure 3). Three types of constraints have usually been considered in channel assignment problem.

i) *The Co-channel Constraint (CCC):* The same channel cannot be simultaneously allocated to a pair of cells unless there is a minimum geographical separation between them. Figure 3

ii) *The Adjacent Channel Constraint (ACC):* Adjacent channels cannot be assigned to adjacent cells simultaneously. In other words, any pair of channels in adjacent cells must have a specified distance. Note that the distance indicates the difference in the channel domain.

iii) *The Co-site Constraint (CSC):* Any pair of channels in the same cell must have a specified distance. This distance for CSC is usually larger than that for ACC.

The goal of channel assignment problem is therefore to find a channel assignment to every call request with the minimum number of channels subject to the above three constraints.

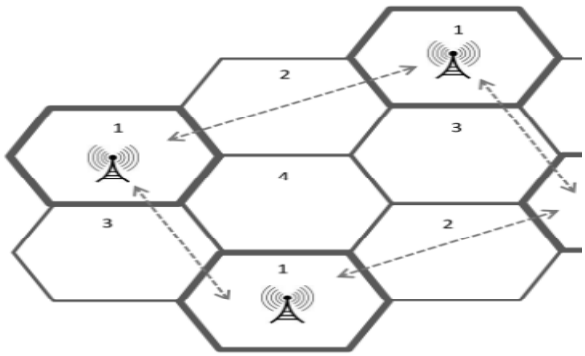


Figure 3: Co-Channel interference (Deepak et al, 2012)

A solution to the CAP must satisfy the hard constraints whereas a soft constraint may be violated. Soft constraints help to maximize the utilization of resources and/or improve the quality of service (Deepak et al, 2012). A main focus of this work is to maximize the utilization of the limited channel resources in order to improve the quality of service provided by the network. Hence in contrast to most works on channel allocation, we will consider together with the hard constraints all soft constraints. The following known soft constraints explained in (Deepak et al, 2012) and (Wu et al., 2010) will be considered:

(a) Packing condition (PC): try to use the minimum number of channels every time a call arrives. This condition encourages the selection of channels already in use in other cells as long as the hard constraints are satisfied. This will minimize the number of channels used by the network and lower the probability of future call blocking

(b) Resonance condition (RC): try to assign the same channels to cells that belong to the same reuse scheme. The purpose of this approach is to leave as many channels as possible to be allocated to other cells belonging to other reuse schemes. Consequently, the probability of causing co-channel interference in the system is reduced.

(c) Limiting rearrangement (LR): try to assign, whenever possible, the same channels assigned before to the existing calls, thus limiting the reassignment of channels. Channel reassignment is the process of transferring an ongoing call to a new channel without call interruption. Such reassignment in the entire cellular network upon the arrival of a new call will obviously result in lower call blocking probability. However, it is complex, both in terms of time and computation (Wu et al., 2010). Therefore, the reassignment processes should be limited to a low level. On this account, limiting rearrangement condition is used to prevent excessive reassignment in a cell (Wu et al., 2010).

Channel assignment in this work will follow the problem formulation by (Funabiki et al, 2000) and also used in ((Deepak et al, 2012), (Omid, 2011), (Omid, 2010)). The three constraints for the channel interference in the N- cell system are described by an N x N symmetric compatibility matrix C. The non-diagonal element cij $(i \neq j)$ of C represents the minimum distance between a channel assigned to cell $\neq i$ and a channel to cell $\neq j$. The diagonal element $C_{ii}$of C represents the distance between a pair of channels assigned to cell $\neq i$. Thus, *CCC* is described by C_{ij}=1, *ACC* is by $C_{ij} \geq 2$ and *CSC* is by $C_{ii} \geq 1$ in C, respectively.

2.4 CHANNEL ALLOCATION AND FREQUENCY REUSE

Any efficient channel allocation scheme must reuse mostly the available frequency by considering the minimum frequency reuse constraints. A channel allocation method always considers important parameters like *frequency reusability, minimum frequency reuse constraints* and *changing traffic*. Frequency reuse enables the support of more users and becomes paramount when considering how the scarce frequency resources can be adequately managed to serve the ever increasing

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

subscribers. (Saghaei et al., 2013), (Arafat et al., 2013) and (Wolfgang et al., 2013). Hence by applying frequency reuse we guarantee more users will be served by the network operator which means more users with same number of frequencies and this leads to decrease in cost. (Jian et al., 2011) proposed a novel technique for dynamic frequency reuse that will lead to more efficient utilization of the shared spectrum, according to them, dynamic reuse of channels will enable a base station to reuse all allocated spectrum. (Hyang Sin et al., 2011) focused on the usage of end-to-end devices as a cellular network to achieve tremendous multimedia load by improving cell capacity by efficiently and intelligently allocating available channels. (Imran et al., 2010) proposed a novel way for self-organizing and adaptive frequency reuse.

3.0 METHODOLOGY

All three hard constraints can be represented with the help of a compatibility matrix C which is N x N symmetric matrix as shown in figure 4, where N is the number of cells in the network. The following can be deduced from the matrix:

- Each diagonal element C_{ij} shows the CSC, which is the minimum separation in frequency between any two channels at cell i.
- If $C_{ij} = 0$; there is no constraint in channel reuse in between cell i and j.
- If $C_{ij} = 1$; there is a CCC
- If $C_{ij} = 2$; there is a ACC
- If $C_{ii} >= 1$; there is a CSC

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & . & . & . & C_{1N} \\ C_{12} & C_{22} & . & . & . & C_{2N} \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ C_{N1} & C_{N2} & . & . & . & C_{NN} \end{pmatrix}$$

Figure 4: Compatibility Matrix

An assignment/allocation channel by cell matrix A (Figure 4.) reflects which channel is allocated to which cell. $A_{i,j}$ is 1 if channel j is assigned to cell i and 0 otherwise .

Also an allocated channel vector V of dimension 1 … number of channels in the network (Figure 5) shows the channels presently allocated to the cell k in which a call has arrived and requesting a channel to service the call request.

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & . & . & . & A_{1,cel} \\ . & & & & & . \\ . & & & & & . \\ . & & & & & . \\ . & & & & & . \\ A_{cha,1} & A_{cha.2} & . & . & . & A_{cha,cel} \end{pmatrix}$$

Figure 4: Allocation/Assignment Matrix

$$V = \begin{pmatrix} \mathbf{V_{k,1}} \\ . \\ . \\ . \\ . \\ \mathbf{V_{k,cha}} \end{pmatrix}$$

Figure 5: Selected channel vector

3. 1 The Evolutionary Neural Network Models

In addition to the hard constraints already considered, our own approach to the problem gives full consideration to both hard constraints and soft constraints. The set of equations (1) to (6) defines our approach to the formulation and representation of the CAP as an optimization problem incorporating all three hard constraints (CCC, CSC and ACC) and soft conditions (PC, RC and LR) already explained. Equations (1) and (2) are similar to the functions used by (Omid, 2011) to represent the CCC, ACC and CSC hard constraints. We have however introduced weights to determine the significance of the hard constraints. Equations (3), (4) and (5) are similar to functions used by (Deepak et al, 2012) to represent the soft conditions. We have introduced equation (6) to the combined functions used separately by (Omid, 2011) and (Deepak et al, 2012) to arrive at equation (7) which is now our new optimization function for the CAP problem in this work.

$$E_{ACC,CCC} = \sum_{j=1}^{cha} \sum_{i=1}^{cel} \sum_{p=1}^{cha} V_{k,j} * A_{i,j} * ACCC_{kjip} \quad (1)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Where

$$ACCC_{kjip} = \begin{cases} 1 \; if \; i \neq k \; and \; C_{ki} > 0 \; and \; j - (C_{ki} - 1) \leq p \leq j + (C_{ki} - 1) \\ 0 \quad otherwise \end{cases}$$

Equation (1) states the adjacent channel constraint and co-channel constraint. Here V_k is an output vector for cell k, with dimension channel (cha). $V_{k,j}$=1 if channel j is selected to be assigned to cell k, otherwise $V_{k,j}$=0. Here k signifies the cell in which call arrives. The energy function $E_{ACC,CCC}$ increases if a channel j which is assigned in cell i is selected by cell k and interference occurs as a result of that selection. It thus ensures that solutions with no interference give betterfitness values. $A_{i,j}$is the ijth element of the assignment table A, which is 1 if channel j is assigned to cell i, and 0 otherwise.

$$E_{CSC} = \sum_{j=1}^{cha} \sum_{i=1}^{cha} V_{k,j} * A_{k,j} * CSC_{ij} \quad (2)$$

Where

$$CSC_{ij} = \begin{cases} 1 \quad if \; i \neq j \; and \; j - (C_{kk} - 1) \leq i \leq j + (C_{kk} - 1) \\ 0 \quad otherwise \end{cases}$$

$$E_{PC} = \sum_{j=1}^{cha} \sum_{i=1}^{cel} V_{k,j} * A_{i,j} * \frac{(1 - CCC_{ik})}{dist_{ik}} \quad (3)$$

i ≠ k

Equation (3) states the *packing condition*. The energy decreases if channel j assigned to cell k is also selected by cell i andCCC_{ik} =0. Energy reduction depends on the distance between i and k. The packing condition requires that a channel, in use in one cell, should be reused in another cell as close as possible without the channels interfering with each other so that the number of channels used by the network is minimal, thereby lowering the probability of future call blocking in other cells. If this condition is satisfied, it further reduces the energy function

$$E_{RC} = \sum_{j=1}^{cha} \sum_{i=1}^{cel} V_{k,j} * A_{i,j} * (1 - RC_{ik}) \quad (4)$$

i ≠ k

Equation (4) symbolizes the *resonance condition*. Where RC_{ik} is a function whose value is 1 if cells i and k belongs to the same reuse scheme, otherwise 0. The resonance condition tries to ensure that same channels are assigned to cells that belong to the same reuse scheme, as far as possible.

$$E_{LR} = \sum_{j=1}^{cha} V_{k,j} * A_{k,j} \quad (5)$$

Equation (5) subtracts 1 from the energy function when a channel j assigned in cell k for an existing call before the arrival of a new call, remain the channel assigned to the call without reassigning new channel for that existing call. Limiting rearrangement try to assign, whenever possible, the same channels assigned before to the existing calls, thus limiting the reassignment of channels. Channel reassignment is the process of transferring an ongoing call to a new channel without call interruption. Such reassignment in the entire cellular network upon the arrival of a new call will obviously result in lower call blocking probability, but it is complex, both in terms of time and computation. Therefore, the reassignment processes should be limited to a low level. On this account, limiting rearrangement condition is used to prevent excessive reassignment in a cell (Wu et al., 2010).

Equation (6) is introduced to discourage channels being allocated to cells that have reached the maximum number of channels (Mcha) specified for each cell in the network. If the number of channels already assigned to a cell requesting for channel has reached Mcha, it will maximize the function value and makes it a lesser choice for selection. That is it will reduce the chances of the channel being selected for allocation to the cell.

$$E_{Mcha} = Mcha - \sum_{j=1}^{cha} V_{k,j} \quad (6)$$

From equations (1) to (6), the energy function E for cell k becomes

$$\sum_{j=1}^{cha} \sum_{i=1}^{cel} \sum_{p=1}^{cha} V_{k,j} * A_{i,j} * ACCC_{kjip} + \sum_{j=1}^{cha} \sum_{i=1}^{cha} V_{k,j} * A_{k,j} * CSC_{ij} - \sum_{j=1}^{cha} \sum_{i=1}^{cel} V_{k,j} * A_{i,j} * \frac{(1 - CCC_{ik})}{dist_{ik}} + \sum_{j=1}^{cha} \sum_{i=1}^{cel} V_{k,j} * A_{i,j} * (1 - RC_{ik}) - \sum_{j=1}^{cha} V_{k,j} * A_{k,j} - (Mcha - \sum_{j=1}^{cha} V_{k,j})$$

$$E = W_1 * E_{ACC,CCC} + W_2 * E_{CSC} - W_3 * E_{PC} + W_4 * E_{RC} - W_5 * E_{LR} - E_{Mcha} \quad (7)$$

The constants W_1, W_2, W_3, W_4 and W_5 are weights introduced to determine the significance of the various conditions

Our task becomes to optimize the energy function represented in equation (7). This energy function can be minimized using Hopfield neural

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

network with the appropriate interconnection weights and the external inputs. However the drawback with the use of neural networks in solving optimization problems is that they produce local optimal values rather than global optimal values and that the solution depends on the initial values of the neurons and weights(Jamal and Zouchair, 2011). Many approaches have been used to try control the converging of the neural network at local minimal considering the challenge of being able to set the initial values of the network rightly to be able to arrive at a global optimal solution.

3.2 EVOLUTIONARY NEURAL NETWORK MODEL

In this work, we will approach the combination of GA and NN in addressing the CAP from three different angles. In each case, the number of genes in each chromosome of the GA is equivalent to the number of neurons in the NN and equal to the number of cells in the network. Each chromosome represents a channel in the network hence the number of channels in the network is equivalent to the number of chromosomes that make up a population in each generation of the GA and determines the stopping criteria in the NN.

In the first approach figure 5, the initial population of the GA is randomly chosen and subjected to the GA operations of selection, crossover and mutation to produce a near optimal solution. The optimal solution from the GA process is then used to initialize the values of the NN and trained to produce an optimal solution.

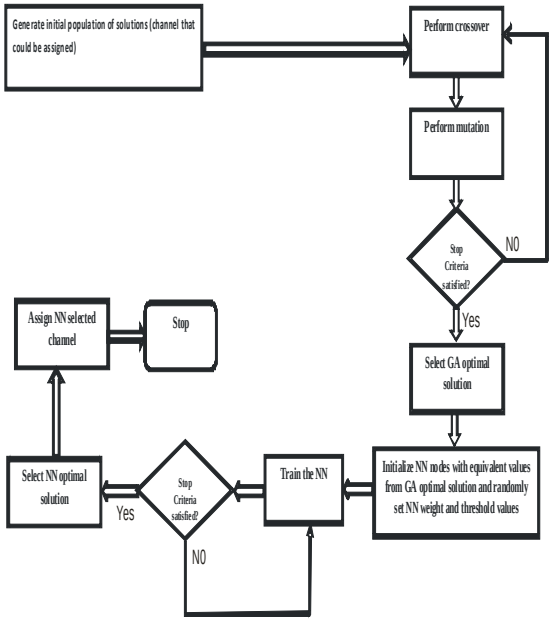


Figure 5: Flow diagram for the first approach

In the second approach Figure 6, each chromosome in the randomly selected initial population of the GA is used to initialize the values of the NN and trained to produce a feasible solution; the optimal solution produced for each chromosome replaces the chromosome in the initial population of the GA. The present population of feasible solutions in the GA is then subjected to the GA operations of selection, crossover and mutation to produce an optimal solution.

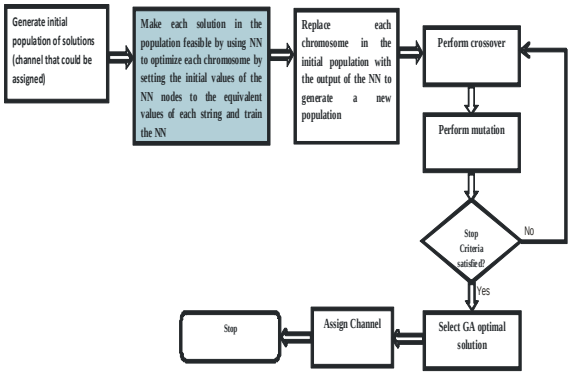


Figure 6: Flow diagram for the second approach

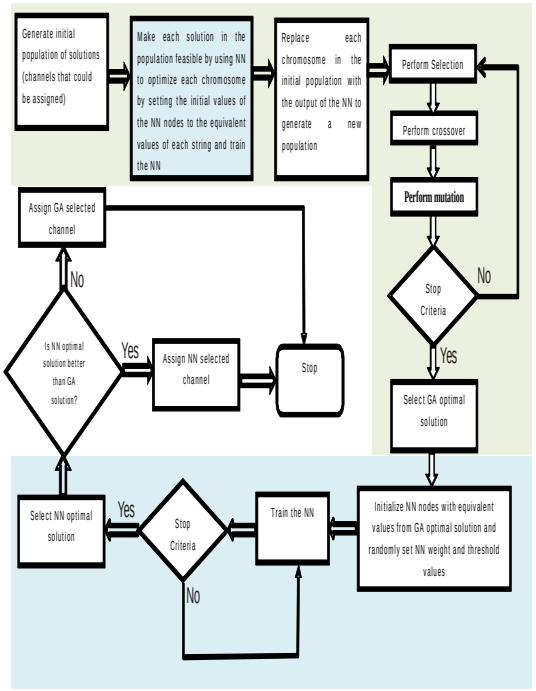


Figure 7: Flow diagram for the third approach

In the third approach figure 7, the optimal solution produced in the GA in the second approach is once again used to initialize the values of the NN and trained to see if a better optimal solution could be obtained. The solution obtained is compared with that generated from the second approach and the better one is chosen.

In the approaches discussed, the combination of GA and NN is such thatone is either used to set the initial values of the other or ultimately used to obtain the final optimal solution. GA will be used to set appropriate initial input data for the NN

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

and NN will be used to initially give local optimization to each chromosome in the GA to make them feasible. Genetic Algorithms are well known to be able to provide very quickly an acceptable near optimal solution (Maximiano et al., 2008). Meaning that genetic algorithms can quickly converge to the field of minimal errors, but it is difficult to achieve the final convergence (Junsheng, 2013). Unlike Neural Networks, Genetic Algorithm ensures a quicker convergence to the near-optimal solution (Rincy and Jude, 2011) which makes it a good technique for selecting the initial values of the neurons of a neural network. It ensures that the neural net converges to an optimal solution since the initial values which determine whether an optimal solution is attained or not would be values that had generated a near optimal solution from the GA. GA have been combined with NN to solve diverse optimization problems in the past (Jusheng, 2013), (Safa and Mohamed, 2013), (Partha et al., 2012), (Deepak and Avinash, 2013)

3.2 PROPOSED ARCHITECTURE

From the system architecture (Figure 8), the network parameters and the fitness function are available to both the genetic algorithm and neural network modules of the system. The output generated by the genetic algorithm module can either be used to initialize the neuron values of the NN or taken as the optimal solution depending on the approach implemented. The neural network module of the system could be used to make the randomly generated solutions of the GA feasible solutions before applying the genetic operations to generate an optimal solution or its solution can be taken as the optimal solution depending on the approach being implemented. In the third approach, output of the genetic algorithm module and that of the neural network module of the system are compared by the solution comparator to decide the best optimal solution

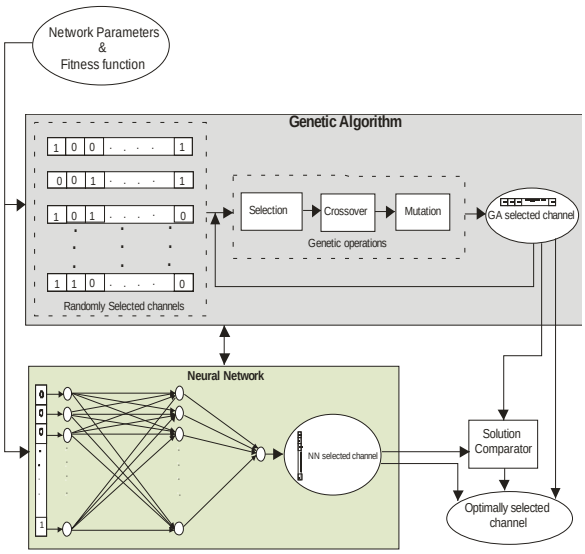


Figure 8: The Proposed System Architecture

3.3 EXPECTED CONTRIBUTION TO KNOWLEDGE

One major contribution of this work is the consideration of all possible constraints (both hard and soft constraints) and elimination of unrealistic assumptions in addressing the CAP. The soft constraints have been proved to add value to the quality of service provided when considered.

Another contribution is the expected reduction in the Call drop and call block probabilities by the combination of two search techniques to solve the CAP. The combination of GA known to produce a near global optimal solution over a large search space and NN well known to easily produce a local optimal solution over a small search space is expected to produce a better optimal solution than existing methods in the allocation of channels which will in turn lead to better results in attempting to reduce the call dropping and blocking probabilities.Approaching the CAP with a combined GA and NN techniques from three different angles will help the understanding of how best the two techniques can be combined to solve the CAP and not just combining them

Ultimately the successful implementation of our proposed model is expected to reduce congestion in mobile communication network following the expected reduction in call blocking and call dropping probabilities across the network.

CONCLUSION

This paper demonstrates the efficiency of integrating genetic algorithm and neural network approaches with dynamic allocation technique in channel assignment for mobile networks to improve the performance and flexibility of channel allocation.

REFERENCES

Aizaz Z., Tirmizi A, and Raeen S. (2012), "Implementation of Genetic Algorithm in Dynamic Channel Allocation".*International Journal of Scientific Engineering and Technology, Vol. 1 issue 2 pg: 108 – 11,1 2012*

Alireza S., Shirazi G., and Amindavar H. (2006), "A hybrid method for channel assignment problemsin cellular radio networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '06)*, vol. 3, pp. 1260–1265, 2006.

Arafat M., Christian B. and Abdullah A. A. (2013)" Dynamic channel allocation in cellular networks" *International journal of Innovative and Applied studies vol. 3 July 2013 Pp 608-615*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Ch´avez-Santiago R., Gigi E., and Lyandres V. (2006), "Channel assignment for cellular mobile networks with non-uniform cells—an improved heuristic algorithm," *IEE Proceedings: Communications*, vol. 153, no. 1, pp. 61–68, 2006.

Chandralekha M. and Praffula K. B., (2010), "Minimization of number of handoff using Genetic Algorithm in heterogenous wireless networks", International Journal of Latest Trends in Computing (E-ISSN: 2045-5364) 24 Volume 1, Issue 2, December 2010

Chia Y.S., Siew Z.W., Kiring A., Yang S.S. and Teo K.T.K. (2011), "Adaptive hybrid channel assignment in wireless mobile network via genetic algorithm", *11th IEEE International Conference on Hybrid Intelligent Systems,* pp. 511-516, 2011.

Chia Y.S., Siew Z.W, Yew H. T, Yang S. S. and TeoK. T.K. (2012), "An Evolutionary Algorithm for Channel Assignment Problem in Wireless Mobile Networks" *ICTACT Journal on Communication Technology, December 2012, Volume: 03, Issue: 04*

David K (2010), "Tutorial demo of Neural Networks and Genetic Algorithms" Masters Thesis 2010.FacultasArtisInformatica, UniversitasMasarykiana

Deepak D. and Avinash W. (2013) "Study of Hybrid Genetic Algorithm Using Artificial Neural Network in Data Mining for the Diagnosis of Stroke Disease" *International Journal of Computational Engineering Research. Vol. 03, Issue, 4April, 2013*

Deepak K. S, Srinivas K. and Bhagwan Das D.(2012), "A Useful Metaheuristic for Dynamic Channel Assignment in Mobile Cellular Systems" *International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 1, Nº 6. 2012*

Dorne, R. and Hao, J., (1995),"An evolutionary approach for frequency assignment in cellular radio networks". IEEE International Conference on Evolutionary Computation, Vol. 2, November (1995) 539-544

Duque-Ant´on, M., Kunz, D. and Ruber, B. (1993), "Channel Assignment for Cellular Radio Using Simulated Annealing". IEEE Transactions on Vehicular Technology, Vol.42, Issue:1, February (1993) 841-855

Eisenblater, A. (1997), "A Frequency Assignment Problem in Cellular Phone Networks (extended abstract). Network Design", Connectivity and Facility Location, Volume 35, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, AMS (1997)

El Alami L. S., Kudoh E., and Adachi F. (2006), "On-demand channel assignment using channel segregation for uplink DS-CDMA multi-hop virtual cellular network," in *Proceedings of the 63rd IEEE Vehicular Technology Conference (VTC '06)*, vol. 2, pp.713–717, July 2006.

Fernando, X.N. and Fapojuwo, A.O. (2002), "A Viterbi-like algorithm with adaptive clustering for channel assignment in cellular radio networks". IEEE Transactions on Vehicular Technology Vol. 51, Issue:1, January (2002) 73-87

Farid T., Ngom A., and Jaekel A. (2007), "Integrated hybrid channel assignment and distributed power control in wireless cellular networks using evolution strategy," in *Proceedings of the IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP '07)*, pp. 293–300, 2007.

Fu, X., Bourgeois, A. G., Fan, P. and Pan, Y., (2006)"Using a genetic algorithm approach to solve the dynamic channel-assignment problem". Int. J. Mobile Communications, Vol. 4, Issue:3 (2006) 333-353.

Funabiki, N., Okutani, N. & Nishikawa, S. (2000),"A Three-Stage Heuristic Combined Neura –Network Algorithm for Channel Assignment in Cellular Mobile Systems". *IEEE Trans. Veh. Technol.*,Vol.49,pp. 397–403

Gamst, A. and Rave, W. (1982), "On frequency assignment in mobile automatic telephone systems."*Procedeeng on Global Communication*. pp. 309-315

Ghosh S. C., Sinha B. P., and Das N. (2006), "Coalesced CAP: an improved technique for frequency assignment in cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, pp. 640–653, 2006.

Gomes, F. C., Pardalos, P., Oliveira, C. S., and Resende, M. G.C. (2001),"Reactive GRASP with path relinking for channel assignment in mobile phone networks" In Proceedings of the 5th international Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications. DIALM '01 (2001) 60-67

Hoydis J., Kobayashi M. and Debbah M. (2011), "Green small-cell networks", *IEEE Veh. Tech. Maga.*, vol.6,no.1, pp.37-43, 2011.

Hyang Sin C., Jaheon G., Bum-Gon C., and Min Young C. , "Radio resource allocation scheme for device-to-device communication in cellular networks using fractional frequency reuse," in *Communications (APCC), 2011 17th Asia-Pacific Conference on*, 2011, pp. 58-62.

ImranA., ImranM.A., andTafazolliR.."A novel Self Organizing framework for adaptive Frequency Reuse and Deployment in future cellular networks," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, 2010, pp. 2354-2359.

Jamal E. and Zouchair G. (2011), "Evolutionary Neural Networks Algorithm for the Dynamic Frequency Assignment Problem" International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 3, June 2011

Jian L., Hui Y., Haokai C., Zhongnian L., Shouyin L. , "A Novel Dynamic Full Frequency Reuse Scheme in OFDMA Cellular Relay Networks," in *Vehicular Technology Conference (VTC Fall), 2011 IEEE*, 2011, pp. 1-5.

Jiming C., Qing Y., Peng C., Youxian S., Yanfei F., and Xuemin S. (2011), "Game Theoretical Approach for Channel Allocation in Wireless Sensor and Actuator Networks", *IEEE transactions on automatic control, vol. 56, no. 10, October 2011*

Junsheng J. (2013), "BP Neural Network Algorithm Optimized by Genetic Algorithm and Its Simulation". IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 2, January 2013

Kanal L. and Kumar V. (1988), "Search in Artificial Intelligence", N.Y., 1988.

Khanbary L.M. O. and Vidyarthi D. P. (2009), "Reliability-Based Channel Allocation Using GeneticAlgorithm in Mobile Computing" *IEEE Trans. Veh.Technol.*,vol. 58, no. 8, pp. 4248–4256, October 2009.

Khanbary L.M.O. and Vidyarthy D.P. (2008), "A GA Based Effective Fault-Tolerant Model for Channel Allocation in Mobile Computing" *IEEE Transaction on Vehicular Technology*, 57(3).

Khanbary L.M.O. and Vidyarthy D.P. (2009), "Channel Allocation in Cellular Network Using Modified Genetic Algorithm*" International Journal of Artificial Intelligence Vol. 3 No 9, 2009*

Khanbary L.M.O. and Vidyarthy D.P. (2009), "A GPS Based Deterministic Channel Allocation for Cellular Network in Mobile Computing*" International Journal of Business Data Communication and Networking Vol. 5, Issue 4, 2009.*

Khandare S M. and Sedamker R. R. (2012)"Efficient Scheme for Dynamic Channel Allocation Using Intelligent Agent in Mobile Communication" *International Conference on Advances in Communication and Computing Technologies (ICACACT) 2012 Proceedings published by International Journal of Computer Applications(IJCA)*

Koutsopoulos I. and Tassiulas L. (2007), "Joint optimal access point selection and channel assignment in wireless networks,"*IEEE/ACMTransactions on Networking*, vol. 15, no. 3, pp. 521–532, 2007

Kunz, D. (1991),"Channel assignment for cellular radio using neural networks".*IEEE Trans. Veh.Technol.*,Vol.40,pp. 188-193

Lai W. K. and George G. C. (1996), "Channel Assignment Through Evolutionary Optimization", *IEEE Transactions on Vehicular Technology*, vol. 45, no. 1, pp. 91-96, February 1996.

Lakshmi M.L.S.N.S., Ratnavathi M.S.L. and Gopi Krishna S. (2012),"An insight to call blocking probabilities of channel assignment schemes" *International Journal of Advances in Engineering & Technology, May 2012*.

Li L., Tao J., and Li F.(2006), "Dynamic channel assignment performance analysis in multiservice hierarchical wireless networks," in *Proceedings of the 1st International Conference on Communications and Networking in China (ChinaCom '06)*, pp. 1–5, 2006.

Li L., Tao J, and Xiaofang T. (2006), "Dynamic channel assignment performance analysis in multiservice hierarchical wireless networks," in *Proceedings of the 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications(PIMRC '06)*, pp. 1–5, September 2006.

Li X. J. and Chong P. H. J. (2007), "A dynamic channel assignment scheme for TDMA-based multihop cellular networks," in *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '07)*, pp. 815–818, September 2007.

Lima M. A. C., Araujo A. F. R., and Cesar A. C. (2007), "Adaptive genetic algorithms for dynamic channel assignment in

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

mobilecellular communication systems," *IEEE Transactions On Vehicular Technology*, vol. 56, no. 5, pp. 2685–2696, 2007.

Lima M. A. C., Araujo A. F. R. and Cesar A. C.(2007), Adaptive genetic algorithms for dynamic channel assignment in mobile cellular communication systems, *IEEE Trans. Veh. Technol.*, vol.56, no.5, pp.2685-2696, 2007.

Liu, X., Pardalos, P. M., Rajasekaran and Resende, M.G.C. (2000),"A GRASP for Frequency Assignment in Mobile Radio Neworks". DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 52 (2000) 195-201

Maximiano M., Vega-Rodriguez M.A., et al., (2008), "Analysis of Parameter Settings for Differential Evolution Algorithm to Solve a Real-World Frequency Assignment Problem in GSM Networks". The Second International Conference on Advanced Engineering Computing and Applications in Sciences (ADVCOMP'08),Valencia, Spain, 2008

Mohapatra S.S., Roy K.K., Benerjee S. and Vidyarthy D.P. (2008) *IEEE transaction on mobile computing*, 57(7).

Mona M.E. and Mohamed E. E. (2008), "Fixed Channel Assignment and Network Capacity Estimation Algorithms for Cellular Networks with heavy Load Conditions" *INFOS2008, March 27-29, 2008 Cairo-Egypt*

Nie J. and Haykin S. (1999), "A Q-learning Dynamic Channel Assignment Technique for Mobile Communications Systems", *IEEE Trans.On Vehicular Technology*, vol. 48, no. 5, pp.1676- 1687, September 1999.

Nie J. and Haykin S. (1999), "A Dynamic Channel Assignment Policy Through Q-Learning", *IEEE Transactions on Neural Networks*, vol. 10, no. 6, pp. 1443-1455, November 1999.

Ojesanmi O. A, Ojesanmi A. and Makinde O. (2009), "Development of Prioritized Handoff scheme for Congestion Control in Multimedia Wireless Network," Proceedings of the World Congress on Engineering, London U.K. Vol I, July 1 - 3, 2009.

Ojesanmi O. A., Ojesanmi S. O. and Ojesanmi A. A., (2010), "Multi-Threshold with One-Level Buffer for Call Admission Control in Mobile Cellular Systems". *International Journal of Electrical and Computer Engineering. ISSN 0974-2190 Volume 2, Number 1 (2010), pp. 63-70*

Omid M. (2011), "A Hopfield Neural Network for Channel Assignment Problem in Cellular Radio Networks" *Computer and information Science Vol 4, No 1, January 2011*

Omid M. (2010), "Fixed Channel assignment and Neural Network Algorithm for Channel Assignment Problem in Cellular Radio Network" *Computer and information Science Vol 3, No 4, November 2010*

Papazoglou P.M., Karras D.A. and Papademetriou R.C. (2006), "High Performance Novel Hybrid DCA algorithms for efficient Channel Allocation in Cellular Communications modeled and evaluated through a Java Simulation System", WSEAS Transactions on Communications (with impact factor), ISSN 1109-2742,Issue 11, vol 5, 2006, pp.2078- 2085

Papazoglou P.M., Karras D.A. and Papademetriou R.C. (2006), "On new dynamic channel assignment schemes and their efficient evaluationthrough a generic simulation system for large scale cellulartelecommunications", HERMIS, An International Journal of ComputerMathematics and its Applications, ISSN 1108-7609, Vol. 6, 2006.

Papazoglou P.M., Karras D.A. and Papademetriou R.C. (2006), "Novel DCAalgorithms for efficient Channel Assignment in CellularCommunications and their evaluation through a generic Java SimulationSystem ", 6th WSEAS Int. Conf. on Simulation, Modeling and Optimization (SMO '2006)

Papazoglou P.M., Karras D.A. and Papademetriou R.C. (2007), "Simulating andEvaluating Dynamic Channel Assignment Schemes in WirelessCommunication Networks through an Improved Multi-Agent System",3rd Indian International Conference on Artificial Intelligence (IICAI-07),2007

Park S. Y. (2013) "Genetic Algorithm-Based Adaptive Channel Allocation for Best-Effort Traffic Users in Microcellular Systems" *International Journal of Innovative Computing, Information and Control*. Volume 9, number 5, May 2013.

Partha P.S., Banshidhar M. and Madhumita P. (2012), "Performance Analysis of Neural Networks Training using Real Coded Genetic Algorithm" *International Journal of Computer*

Applications (0975 – 8887) Volume 51– No.18, August 2012

Patra S. S. M., Roy K., Banerjee S. and Vidyarthi D. P. (2006), "Improved genetic algorithm for channel allocation with channel borrowing in mobile computing", *IEEE Trans. Mobile Computing*, vol.5, no.7, pp.884-892, 2006.

Pinagapany S. and Kulkarni A. V. (2008), "Solving channel allocation problem in cellular radio networks using genetic algorithm", *Proc. of IEEE COMSWARE*, pp.239-244, 2008.

Prajapati A. K., Ghosh R. K., and Mohanty H. (2006), "A self-adaptive hybrid channel assignment scheme for wireless communication systems," in *Proceedings of the 9th International Conference on Information Technology (ICIT '06)*, pp. 94–95, December 2006

Ramesh B. T. K. and Jana N. D. (2012) "An Optimized way for Static Channel Allocation in Mobile Networks Using Genetic Algorithms.*International Journal of Computer Application (0975 – 8887) Volume 45 – No. 19, May 2012*

Rincy G. and Jude H. (2011), "Application of an Optimization Algorithm for Channel Assignment in Mobile Communication" *IJCA Special Issue on "Novel Aspects of Digital Imaging Applications" DIA, 2011*

Safa S. I. and Mohamed A. B. (2013) "Interpretation Trained Neural Networks Based on Genetic Algorithms*" International Journal of Artificial Intelligence & Applications (IJAIA), Vol.4, No.1, January 2013*

Safwat A. M. (2006), "Distributed connection admission control and dynamic channel allocation in ad hoc-cellular networks," in *Proceedings of the International Conference on Digital Telecommunications (ICDT '06)*, p. 52, August 2006.

Saghaei H., Saghaei H and Darvish M. (2013) "A new approach for wireless cellular network design" *Int. J. Communication, network and system services, 2013, vol. 6, 82 -87*

Sancho S., Jose A., Emilio G., Angel M., Christopher T.,(2008),"Optimal switch location in mobile communication networks using hybrid genetic algorithms", Elsevier: Applied soft computing, 2008

Shashi B and Svetlana G. A.A(2012)"Ga Based Model for Distributed Dynamic Fault Tolerance Channel Allocation in Cellular Networks" Journal of Information and Operations Management Volume 3, Issue 1, 2012, pp-316-318

Shyamahe T. and Rahim T. (2004), "Efficient Call Admission Control and Scheduling for GPRS using Genetic Algorithms".*IEEE 2004 Pgs 2528 - 2533*

Sivarajan, K. N. ,McEliece, R. J., & Ketchum, J. W. (1989), "Channel assignment in cellular radio." *IEEE Vehicle Technology Society Conference*. pp. 846-850

Smriti D. (2013) "Optimal Channel Allocation with Hot –Spot Technique in Wireless Network" *International Journal of Computer Application (0975 – 8887) Volume 61 – No. 10, January 2013*

Susil K. S. and Prafulla K. B (2011) "Path Loss-A Parameter That Affects Channel Performance in Mobile Communication" *National Journal of Computer Science & Technology,* vol. 3, pp. 34-36

Susil K. S (2012), "Fixed Channel Allocation in Mobile Network: A Critical Review" *International Research Journal of Research Analysis & Evaluation,* vol. 4, pp. 44-47.

Susil K. S. and Prafulla K. B (2012), "A Comparative Analysis on Channel Allocation Schemes towards an Efficient Assignment Strategy in Cellular Mobile Network" *National Journal of Computer Science & Technology*, vol. 4, pp. 47-51.

Susil K. S. and Prafulla K. B. (2013) "An Efficient Hybrid Channel Allocation Mechanism in Mobile Network" *International Journal of Research in Engineering & Applied Sciences Volume 3, Issue 2 February 2013.*

VidyarthiG., Alioune N. and Ivan S. (2005), "A Hybrid Channel Assignment approach Using an Efficient Evolutionary Strategy in Wireless Mobile Networks" IEEE Transactions On Vehicular Technology, Vol. 54, No. 5, September 2005

Vidyarthi G., Ngom A., and Stojmenovi´c I. (2005), "A hybrid channel assignment approach using an efficient evolutionary strategy in wireless mobile networks," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 5, pp. 1887–1895, 2005.

Vidyarthi G., Ngom A., and Stojmenovi´C I. (2006), "Evolutionary methods in wireless mobile computing," in *Combinatorial Optimization in Communication Networks*, D.-Z. Du, M. Cheng, and Y. Li, Eds., vol. 18, pp. 33–79, Springer, New York, NY, USA, 2006.

Wang, W. and Rushforth, C.K.(1996)"An Adaptive Local-Search Procedure for the

Channel-Assignment Problem (CAP)", IEEE Transactions on Vehicular Technology, Vol.45, Issue:3, august (1996) 459-466

Wolfgang M., Michael G., Wolfgang Z. and Gerhard F. "Interference Mitigation framework for cellular mobile radio network" *International Journal of Antennas and propagation volume 2013, 15 pages.*

Wu X, Jackel A. and Bari A. (2011) "Optimal Channel Allocation with Dynamic Power Control in Cellular Networks" *International Journal of Computer Networks & Communications (IJCNC) Vol.3, No.2, March 2011*

Wu X., Jaekel A., Bari A., and Ngom A. (2010), "Optimized hybrid resource allocation in wireless cellular networks with and without channel reassignment," *Journal of Computer Systems, Networks and Communications Vol. 2010*

Segmentation of Kannada Handwritten Text Line through Computation of Variance

Sunanda Dixit[#1], Suresh Hosahalli Narayan[#2]
[1]*Information Science and Engineering Department*
Dayananda Sagar College of Engineering
Bangalore, India
[2]*Department of Instrumentation Technology*
Bangalore Institute of Technology
Bangalore, India

***Abstract*— Handwritten text line segmentation is an important task of Optical Character Recognition. The proposal discusses a novel technique for Segmentation of Lines for the handwritten text document written in Kannada Language. The algorithm imbibes the approach of finding the Components, bounding box and computing the coefficient of variance. Preprocessing steps such as correction of skewness due to improper scanning as well as removal of noise is assumed to have been performed on the data. Our technique of line segmentation addresses the key complexity issue of variations in the gaps between and across lines as well as curly and skew character text of handwritten document. The purpose of this paper is to segment line by connecting the centroids within the bounding box and employing the coefficient of variance. The method has been tested on unconstrained Kannada scripts. An experimental result obtained by executing the proposed method of line segmentation has achieved high degree of accuracy and high performance.**

***Keywords*— Centroids, Coefficient of Variance, Document Image Processing, Optical Character Recognition, Text line segmentation**

I. Introduction

Segmentation of text line and words in document image processing is a complex and crucial step towards recognition of unconstrained handwritten document. Offline handwritten text segmentation is an important pre-requisite of Optical Character Recognition (OCR). Segmentation errors reduce the accuracy of OCR. The conventional line extraction methods are classified as projection analysis, grouping mechanism, smearing and Hough methodology [1]. Projection based approach works quiet well for machine printed and scanned documents however it may not be suitable for handwritten documents. Piece wise projection profile is suitable for skew angle variation. For constant skew angle text lines global projection profile based methods are very effective. Grouping approaches are effective in handling complex layouts [2]. In case of touching text lines and varying skewness, the accuracy is quiet less. Intensive research effort in the area of line segmentation for textual handwritten documents exists for global languages English, Asian languages Chinese and Japanese. As per the literature most of the handwritten segmentation work has been carried out in Indian Languages like Bangla, Hindi and Gurumukhi. [7][8][9]. Few works have been proposed on south Indian languages like Kannada and Tamil [11] [12].

The proposed technique of line-segmentation of Kannada handwritten document majorly defines two phase. First phase involves classifying the image text into components. The next phase uses the coefficient of variance value for distinguishing and organizing the components into lines. The paper is categorically structured into sections. The work related to line segmentation of text documents are

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

captured in Section II. A brief about Kannada language is provided in section III. The core technique for text line segmentation depicting the method proposed and the dataset used for experimenting as well as Result analysis is presented in Section IV. The conclusion and summarization for work is described in Section V.

II. Related work

The process for segmentation of Text lines from documents that are machine printed is relatively simple and many techniques exist. The challenge is for segmenting the documents that are handwritten. In case of handwritten text documents the problem is diverse and is complicated solely by the nature of writing using hands. So far many research works carried out in text line segmentation in English, Chinese, Arabic and Devanagari [3] [4] [5]. Zhang and Sang [6] proposed tensor voting algorithm for Chinese handwritten text line segmentation. In this method morphological processing is used for generating connected component. Removal of outliners and identification of centroids of the connected component are carried-out using 2-D tensor voting principle. The direction of the vectors as well as the saliency values depicts information for tensors which can be used for segmentation. Gupta et al proposed [10] a methodology for creating line segmentation of a text document taking into consideration the movement of a pen tip. The process involves identifying all the connecting components and locating their centroids. The centriodal values so found are then linked based on the rules of pen tip algorithm. 2D tensor voting algorithm helps to remove outlines. The algorithm also uses the sparse data points to derive perpetual structures namely junctions and curves. The above mentioned techniques work better for text lines that are consistent and parallel in a specified direction. These techniques have a limitation where the text lines are structured inconsistently and lines oriented in different directions.

III. Characteristics of Kannada Script

In this section, we describe briefly the core characteristics in adherence to Kannada script so as to address the difficulties of segmentation. Kannada also known as Canarese is popular in South India. It is widely spoken and declared as official language in the state of Karnataka. Kannada is an ancient language and is derived from Dravidian scripts. The population speaking the language of Kannada amounts to 60 million particularly in the Indian states of Karnataka. The main difficulties for segmenting Kannada scripts are due to complexity of characters. The characters can be any one of the type, standalone vowel, a standalone consonant and a consonant modified by a vowel or one or more consonant and a vowel. There are 16 vowels and 36 consonants including conjunct components (subscript/vatthu) as shown in Figure 1 and Figure 2 respectively .

Fig. 1 Vowels and vowel diacritics with ka

Fig. 2 Consonants

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IV. Proposed Methodology

The block diagram in Figure 3 represents the algorithmic flow of the proposed line segmentation technique for text documentation. The stages are clearly pipelined. The different process blocks are: (a) Componentization Process, (b) Filtering Process and (c) Line Construction Process. The sub-section explains the process flow across each stage.

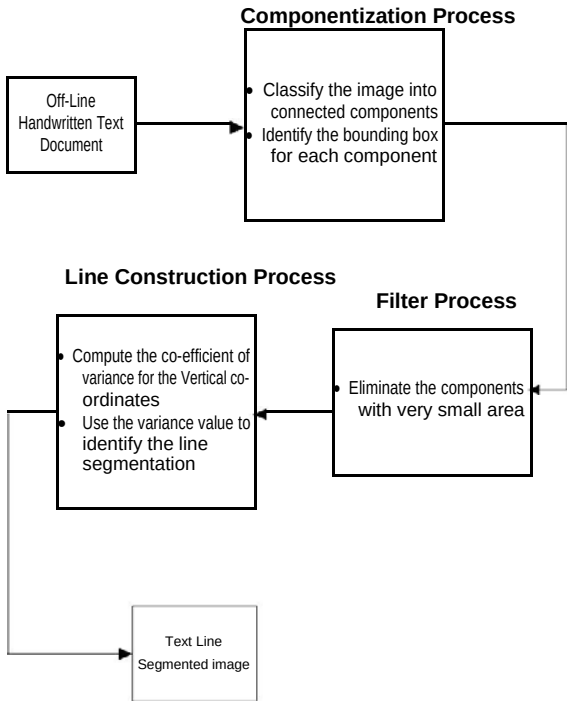


Fig. 3 Block Diagram of the proposed algorithm

The algorithm accepts the raw binary image as the input and identifies all the connected components through the application of two pass algorithm. For each of the components so found, compute the bounding box. Then the filtering logic is applied by calculating the mean area of the entire component and eliminating those whose areas are less than 10 percent of the mean area.

$$\text{Mean Area} = \sum_{i=1}^{N} \frac{Xi * Yi}{N} \quad (1)$$

Where

Xi = Width of the Bounding Box

Yi = Height of the Bounding Box

N = Number of Bounding Box

Threshold Area = Mean Area * 10%

The centroid of all the bounding boxes are then determined. Subsequent to the filteration process, the coefficient of variance for all the connected components along y-axis is computed using the following formula

$$\mu_{Height} = \sum_{i=1}^{N} \frac{Yi}{N}$$

$$\sigma_{Height} = \frac{1}{N} \sum_{i=1}^{N} \left(Yi - \mu_{Height} \right)^2 \quad (2)(3)$$

Where

Yi = Height of the Bounding Box

N = Number of Bounding Box

$$\text{Coefficient of Variance} = \frac{\sigma_{Height}}{\mu_{Height}} \quad (4)$$

Where

σ_{Height} = Standard Deviation of the Component

μ_{Height} = Arithematic Mean of the Component Height

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The formula for finding the variance between two components

$$\text{Variance} = \frac{Yp - Yi}{\mu \text{Height}} \quad (5)$$

Where

Yp = Height of the Parent Component

Yi = Height of the Current Component

μHeight = Arithmetic Mean of the Component Height

The line generation algorithm works based on the following rules Sort the components taking into account the vertical position of the centroid values.

First component is always the top most and is stacked for the line component.

Compute the variance between the top component in the stack with the next component in the list.

If the variance is less than the coefficient of variance value, add the component to the line stack.

Repeat steps 1 to 4 to finalize the image document into line segmentation.

The following formula is used to compute the accuracy of the results

$$\text{Segmentation Accuracy} = \frac{\text{No of Correctly Segmented Lines}}{\text{Total Number of Lines}} \quad (6)$$

Experimental Results

All experiments are conducted on Kannada dataset. This database constitutes a collection of various images of unconstrained handwritten documentation. These text documents are scanned to produce high clarity with a resolution of 300 dpi and saved as binary images. Pre-processing like noise removal and binarization is assumed to have been performed on the data. The dataset used for experiment contains 200 lines of handwritten documents which are collected from different people of various age groups. As per the analysis the number of correctly segmented lines are 192.

A sample input is shown in Figure 4 and the corresponding segmented output is depicted in Figure 5.

The summary of the experimental result is given below

TABLE I

Accuracy of Segmentation

Language	Total Lines	Lines correctly segmented	%of accuracy
Kannada	200	192	96%

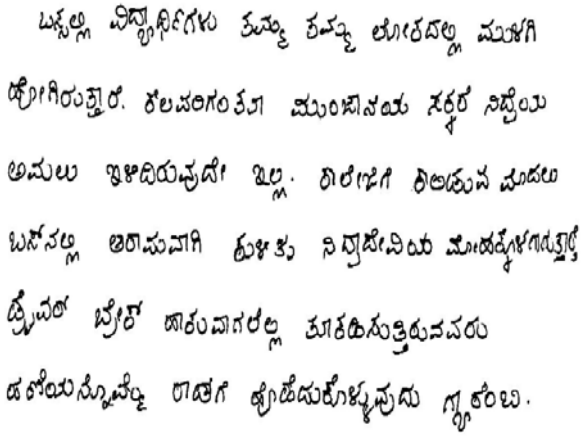


Fig. 4 Input image

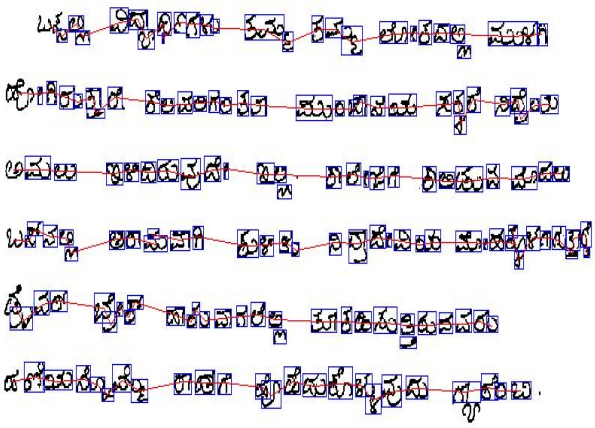


Fig. 5 Output image

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE II

Comparative Study

AUTHOR AUTHOR	SEGMENTATION METHOD	SIZE OF DATASET	SEGMENTATION RATE
Alirez Alaei et al. [14]	Potential Piecewise Separation Line	204	94.98%
Mamatha HR et al. [13]	Morphological operations and projection profile based approach	100	94.5%
Proposed	Computation of Covariance	200	96%

V. CONCLUSION

The paper proposes technique for segmentation scheme for documents that are handwritten in Kannada Language scripts. The analysis of the results of line segmentation clearly indicates that the proposed coefficient of variance method is very robust. The accuracy of the line segmentation algorithm for handwritten document with Kannada language is 96%.

REFERENCES

[1] ZaidiRazak, KhansaZulkiflee, Mohd Yamani IdnaIdris, EmranMohd Tamil, MohdNoorzaily, Mohamed Noor, RosliSalleh, MohdYaakob, ZulkifliMohdYusof, and MashkuriYaacob, "Off-line Handwriting Text Line Segmentation: A Review", International Journal of Computer Science and Network Security, Vol.8, pp. 12-20, 2008.

[2] S. Nicolas, T. Paquet, L. Heutte, "Text line segmentation in handwritten document using a production system", in: Proceedings of International Workshop on Frontiers in Handwriting Recognition (IWFHR'04), pp.245-250, 2004.

[3] F. Yin and C.L.Liu, "A Variational Bayes Method for Handwritten Text Line Segmentation", In the proceedings of 10th International Conference on Document Analysis and Recognition [ICDAR], pp. 436-440, 2009.

[4] T. Sari and M. Sellami, "Overview of Some Algorithms of Off-Line Arabic Handwriting Segmentation", In the proceedings of the International Arab Journal of Information Technology, Vol. 4, pp. 289-300, 2007.

[5] B.Gatos, A.Antonacopoulos and N.Stamatopoulos, "ICDAR2009 Handwriting Segmentation Contest", In the proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR), pp. 1393-1397, 2009.

[6] C. Zhang and G.S. Lee, "Text Line Segmentation in Chinese Handwritten Text Images", In the proceedings of 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp. 1- 3, 2011.

[7] R.Kumar and A.Singh, "Detection and Segmentation of Lines and Words in Gurmukhi Handwritten Text", In the proceedings of Advanced Computing Conference (IACC), pp. 353-356, 2010.

[8] U. Pal and S.Datta, "Segmentation of Bangla unconstrained handwritten text", In the proceedings of 7th International conference on Document Analysis and Recognition, pp. 1128 – 1132, 2003.

[9] N. Kumar Garg, L Kaur and M. K. Jindal "Segmentation of Handwritten Hindi Text", International Journal of Computer Applications , Volume 1 – No. 4, 2010.

[10] J.D.Gupta and B.Chanda, "A Model Based Text Line Segmentation Method for Off-line handwritten Document", In the proceedings of the 12th International conference on Frontiers in Handwriting Recognition , pp. 125 – 129, 2010.

[11] Mamatha H and Srikantamurthy K "Skew Detection, Correction and Segmentation of Handwritten Kannada Document", International Journal of Advanced Science and Technology Vol. 48, 2012.

[12] J.Venkatesh and C. Sureshkumar "Tamil Handwritten Character Recognition Using Kohonon's Self Organizing Map", International Journal of Computer Science and Network Security, VOL.9, 2009.

[13] Mamatha H R and Srikantamurthy K, "Morphological Operations and Projection Profiles based Segmentation of Handwritten Kannada Document", International Journal of Applied Information Systems, Vol. 4, No. 5, pp. 13-19, 2012.

[14] Alireza Alaei, P. Nagabhushan and Umapada Pal "A New Dataset of Persian Handwritten Documents and its Segmentation", IEEE, 2011.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Incremental Classification using Feature Tree

N N Vadnere[#1], R G Mehta[*2], D P Rana[*3], N J Mistry[*], M M Raghuwanshi[**]

[#] *M.Tech, Computer Science Department, Manipal University Jaipur, Jaipur, India*
[1]nishantvadnere@yahoo.com
[*] *SVNIT, Surat, India*
[2]rgm@coed.svnit.ac.in
[3]dpr@coed.svnit.ac.in
[**]*RGCET, Nagpur, India*

Abstract— **In recent years, stream data have become an immensely growing area of research for the database, computer science and data mining communities. Stream data is an ordered sequence of instances. In many applications of data stream mining data can be read only once or a small number of times using limited computing and storage capabilities. Some of the issues occurred in classifying stream data that have significant impact in algorithm development are size of database, online streaming, high dimensionality and concept drift. The concept drift occurs when the properties of the historical data and target variable change over time abruptly in such a case that the predictions will become inaccurate as time passes. In this paper the framework of incremental classification is proposed to solve the issues for the classification of stream data. The Trie structure based incremental feature tree, Trie structure based incremental FP (Frequent Pattern) growth tree and tree based incremental classification algorithm are introduced in the proposed framework.**

Keywords— **Stream Data, Trie, Discretization, Incremental Classification, Feature-Tree**

I. INTRODUCTION

In recent years, stream data have become a rising significant area of research for the database, computer science and data mining communities. Stream data is an ordered sequence of instances that in many applications of data stream mining can be read only once or a small number of time using limited computing and storage capabilities. The classification results change over time when concept drift occurs. This leads to change in data patterns. It is also known as data stream evolution. It has some other issues such as size of data, dimensionality and online streaming. Generally the data mining applications involve millions or billions of records with high dimensionality, which make the task of classification computationally very expensive. Recent applications of Data Mining deals with incremental data so whenever the database is updated the reconstruction of the classification model is needed, which involves many complex, time consuming tasks for the entire dataset. The incremental approach in data mining deals with the model preparation and maintenance incrementally for the updated datasets.

In this paper the framework is proposed for incremental classification for high dimensional stream data. As part of process, first step is to prepare the data which include preprocessing techniques like normalization, missing value replacement, transformation, and discretization. The next step is to make feature tree from preprocessed data. The feature tree is proposed using Trie structure to store the detail of the dataset then FP growth tree is created from the feature tree to reduce the feature for the classification process. The tree based dynamic incremental classification is proposed, which keeps on updating to incorporate the concept drift.

The paper is structured as follows. A brief review of related work is given in Section II which includes preprocessing, data mining algorithm for Trie structure and classification. Section III presents proposed framework and finally Section IV contains summary and future scope.

II. RELATED WORK

To achieve the incremental classification here this section is describing the literature review required at various steps of the proposed framework like data preprocessing, Trie structure and classification.

A. Data Preprocessing

Data preprocessing is the substantial step in data mining. Data mining mainly involve missing value replacement, transformation, normalization, and discretization. The result of data preprocessing is the final training set.

There are many originators of missing value such as broken sensor, erroneous or missing data entries and in some case some attributes make no sense for some type of objects. It is necessary to replace the missing value otherwise the analysis could lead to meaningless denouement. The clear

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

way is to replace the missing value by mean value in case of numeric attributes and modus in case of categorical value [1]. S. McClean et al. proposed a technique to replace the missing value by making rules based on background knowledge [2]. J. J. Shen et al. recommended Rule Recycle bin technique which restate and construct the rules in order to get more complete attributes value association rule. This will enable the database recovered to advance the accuracy and completion rate as well as gain the validity of missing value completion [3]. M. Shyu et al. designed a framework named F-DCS for replacing missing value which accepts the basic concept of conditional probability theories. This framework can manage both nominal and numeric values with a high degree of accuracy when it is compared with other techniques such as using minimum, average and maximum value [4]. R. Malaryizhi et al. enlisted K-NN classifier that performs better than K-means clustering in missing value imputation [5]. N. Devi et al. projected to replace the missing value by mean and median of clusters and got the more accurate result for their classification analysis for different number of clusters [6].

Normalization is also important task in data preprocessing to reduce unwanted variation either within or between arrays. Normally normalization can be done on data with three ways such as Z-Score, by decimal scaling or min-max normalization. There are main two types of normalization based on a) distance and b) proportion. Distance based normalization includes vector based that is on Euclidian distance and linear based normalization which can correct skewedness in data. Proportion normalization includes non-monotonic normalization which is on Z-Score. The normalization property requires that the range of a sameness or distance measure lies within a fixed range.

Transformation is also a valuable step in data preprocessing. Transformation almost compresses the maximum data. Transformation mainly involves smoothing, aggregation, generalization and discretization. A. Kusiak et al. introduced new transformation method named feature bundling [7]. When this transformation technique applied to a training data set it embellish classification accuracy of the decision rules generated from this set. Although bundling is destined for integer, categorical and normative features, it can be continued with continues value, for example by using regression function.

Discretization is a process of transforming continuous attribute value into finite set of intervals to generate attributes with a smaller number of distinct values. There are many types of discretization methods such as Direct vs. Incremental, Single vs. Multi attribute, Supervised vs. Unsupervised, Bottom up vs. Top down. For discretization CAIM algorithm is popular [8]. But one of the major drawbacks of CAIM is its stopping criterion, depends on the number of target classes. When the number of target classes is large, its performance drops. C. J. Tsai et al. invented a new discretization algorithm depends on class attribute Contingency Coefficient [9]. They have shown the outperforming results in term of classification accuracy. W. Qu et al. proposed a new Chi2 algorithm named Rectified chi2 algorithm [10], which looks a new merging standard as the basis of interval merging and discretizes the real value attributes absolutely and rationally. J. Ge et al. presented new discretization algorithm for uncertain data which employ both the formula based and sample based probability distribution function [11]. Results showed that algorithm can help the Naive Bayesian Classifier to extent higher classification accuracy. Q. Zhu et al. proposed a novel and effective supervised discretization algorithm based on correlation maximization by using multiple correspondence analysis which is a technique to catch the correlation between feature intervals and classes [12]. The algorithm naturally generates a better set of feature intervals by maximizing their correlations with the classes and gains the classification performance. K. Sriwanna et al. fabricated the Enhanced CAIM (ECAIM) which is the extended version of CAIM [13]. ECAIM is proposed by two modifications. First modification is extended from CAIM to turn in to a real increment discretization method by upgrading the stopping criterion. The second modification is the multi attribute methods by simultaneously considering all attributes instead of a single attribute. Results demonstrate the ECAIM performs better in both synthetic and real world datasets. R. G. Mehta et al. proposed a Modified CAIM (MCAIM) because the results of CAIM are not sufficient in some cases [14]. Intervals generated by MCAIM discretization are more in numbers so some intervals are needed to merge without loss of discretization information. For merging they used CAIR criterion. Experiments showed that MCAIM with merging discretization gives improved results than CAIM and MCAIM without merging discretization methods.

B. Data Mining Algorithms using Trie Structure

Trie is an ordered tree data structure that is used to store associative array or dynamic set where the keys are usually strings and the term comes from retrieval [15]. Trie was first proposed by R. Braindais [16]. The most common use of Trie is to represent a set of string [17], [18], [19] which is used in dictionary management. The Trie is also used in different areas such as text compression [20], natural language processing [21], [22], searching for reserved words for a

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

compiler [23], pattern matching [24], [25], IP routing tables [26].

In data mining, visual representation can comfort in strengthening the ability of analysing, understanding the techniques, patterns and their assimilation. But for incremental tree the detail visualization process are rarely introduced. Z. Abdullah et al. have explained the visualization process of constructing the incremental Disorder Support Trie Itemset (DOSTrieIT) data structure from the flat-file dataset. DOSTrieIT is also used as a shrink source of information for building FP Tree [27]. Later year the same author presented a scalable technique to discover items support from Trie data structure [28]. Generally Trie data structure represent frequent patterns via frequent pattern tree (FP Tree). Before the FP Tree can be constructed, there are mainly two scanning process involved in the original database. One of them is to determine the item count (item and their count) that fulfils minimum count threshold by scanning the entire database. However if some changes occurred in the database the process must have find the items and their count again so, Z. Abdullah et al. presented a technique called Fast Determination of Item Support Technique (F-DIST) to capture the items counts from their proposed Disorder Support Trie Itemset (DOSTrieIT) [27] data structure. Experiment said that the computational time to capture the items count using F-DIST from DOSTrieIT is importantly outperformed the FP Tree technique about 3 orders of magnitude for 3 UCI benchmark datasets.

C. Classification

Classification maps (or classifies) a data item into one of several predefined categorical classes. Decision tree (DT) can be regarded as a powerful and popular tool for classification and prediction process [29]. DT can be considered more interpretable compared to neural network and support vector machines since they combine more data in an easily understandable format. DT is computationally cheap, easy to use and can deal with uncertainties. It also provides objective analysis to decision making. The drawback of DT is that the whole process requires quantitative data to determine the accuracy of the input.

ID3 (Iterative Dichotomiser 3) is a very simple classification algorithm and based on concept learning ID3, uses information gain to decide which attribute goes into a decision node[30]. The performance of the algorithm degrades as missing values are classified incorrectly. Many researchers proposed modification in the ID3 [31], [32], [33]. C4.5 is the extension of the ID3 but can be used for unavailable values, continuous attribute value range, pruning of the decision tree and rule derivation [34]. The C4.5 generates the decision tree using the information entropy for the classification. Quinlan proposed the C5.0 algorithm which provide improvement over the C4.5. C5.0 is significantly faster, uses memory more efficiently and generates the smaller decision tree than C4.5 [35]. C5.0 supports for boosting which improves the tree and more accuracy. The concept of winnowing is used by C5.0 to automatically winnow the attributes to remove those that may be unhelpful [36].

Schlimmer et al. [37] proposed incremental ID3 algorithm using brute-force method. The drawback of this method is when new data is arrived, an entirely new tree is created. The same author [37] proposed incremental ID4 algorithm. The drawback of this algorithm is that it discards sub trees when new test is chosen for a node. The author Utgoff P.E [38] proposed incremental ID5 which did not discard the sub tree but also not provide guarantee for producing same tree as ID3. The same author [39] Proposed ID5R algorithm. It produces same tree as ID3. It uses recursively updating method for tree's sub nodes. The drawback is that it does not handle numeric values, multicast classification task or missing value. Domingos, Hultern G [40] proposed VFDT (Very Fast Decision Trees) learner reduces the training time for large dataset using subsampling method to incoming data stream. Spencer et al. [41] proposed CVFDT (Concept-adapting Very Fast Decision Trees learner) which can work with concept drift using the sliding window concept on incoming data. The drawback is that it forgot the old data outside the window. G. Fernandes [42] proposed VFDTc which was extension of VFDT for continuous data and concept drift.

III. PROPOSED FRAMEWORK

The stream data is huge in size and with very high dimensionality. As existing data mining algorithms are not suitable for these datasets, the framework for incremental classification is proposed in this paper as shown in the following Fig. 1. The proposed framework contains incremental tree based classification, where the classification model is required to be updated for every unclassified and wrongly classified instances of test and application phase data.

The data containing noise and missing values are to be replaced with suitable data, so the initial stage of the framework deals with preprocessing methods such as missing value replacement, normalization and transformation. The missing value replacement algorithm replaces by mean, median or standard deviation based on data characteristics.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The tree based classification needs categorical data and Trie structured feature tree also deals with categorical data. Discretization is an important stage of the proposed frame work. The preprocessed data is applied to the discretization phase. In this paper MCAIM [14] algorithm is used for discretization which gives better result than previous CAIM algorithm [8]. The prepared data is applied to the feature tree creation phase.

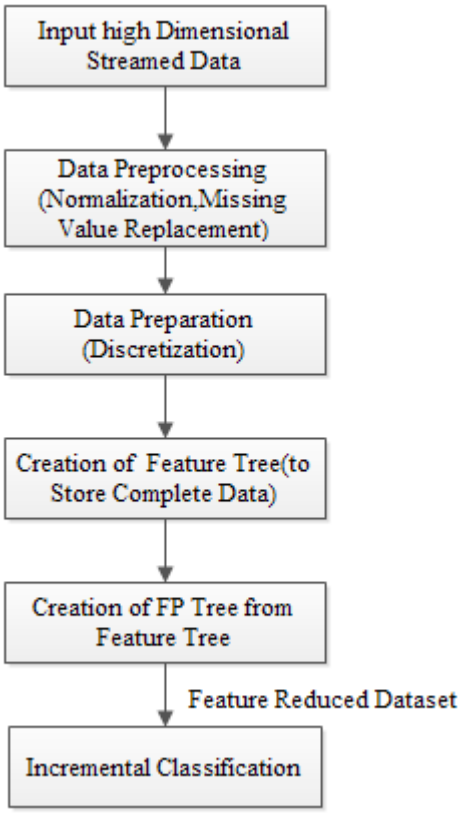


Fig. 1 Block diagram of proposed framework

In classification, as one of the major issue to be dealt with is concept drift. The classification tree can be reconstructed completely in this situation or only affected branch of the tree will be updated. The proposed framework uses updation of the classification tree in this situation, as complete reconstruction will be very time consuming. The storage of the complete data will be a major issue. The Incremental Feature tree (F Tree) is proposed to store the complete historical dataset. Each level of the Trie represents one attribute of the dataset, along with count of every value of the attribute. The complexity of the complete Trie structure will remain same inspite of the varying sized input data.

To improve the efficiency of the classification, the features are required to be reduced as it is a known fact that model tree will not contain every features. Feature reduction phased of the proposed framework uses FP growth algorithm to reduce the features. Trie structured FP tree is used to support incremental FP growth algorithm. The reduced featured dataset is applied to the incremental classification phase. For incremental classification M. Lad et al. proposed ID3 algorithm with combination of CAIM and CAIR for attribute selection criteria [43]. This algorithm is to be updated to update the subtree of the classification tree. The proposed framework allows the classification model to be updated for every unclassified and wrongly classified instances of test and application phase data.

In CVFDT [41] classification algorithm, the window of recent data is managed based on which the classification tree will be updated in case of concept drift. As storage is a major issue for the dataset, it restricts the size of window. The proposed model uses Trie structured feature tree that will reduce this restriction up to remarkable level. The analysis of the classification model, updation in the proposed databases is expected to be performed on the background will not affect the response time of the system. The front end classification tree will be replaced with the background tree in case of updation.

IV. Summary and Future scope

Stream data has several issues such as high dimensionality, size of the database, and online streaming. The proposed framework makes use of Trie structured feature tree to store the dataset with reduced complexity to maintain the online streamed data. The incremental FP Growth algorithm, in support with Novel incremental classification algorithm will support the online response for the classification process without sacrificing the classification accuracy.

Acknowledgement

We would like to take this opportunity to convey our sincere gratitude to all those intellectuals concerned for their magnanimous vision by virtue of which we have been guided through to accomplish our mission. We express our heartiest gratitude to Prof. R. G. Mehta and Prof. D. P. Rana for their valuable guidance, moral support and believing in us.

References

[1] Cw.flek.cvut.cz/lib/exe/fetch.php/cources/ac4m33sad/2_tutorial.pdf.

[2] S. McClean, B. Scotney and M. Shapcott, “Using Background Knowledge with Attribute-Oriented Data Mining” Knowledge Discovery and Data mining (Digest no, 1998/310), IEE colloquiumon, 1998, pp. 1/1-1/4.

[3] J. Shena and M. Chen, "A Recycle Technique of Association Rule for Missing Value Completion" in Proc. AINA'03, 2003, pp. 526-529.

[4] M. Shyu, I. P. Appuhamilage, S. Chen and L. Chang, "Handling Missing Values via Decomposition of the Conditioned Set", IEEE Systems, Man, and cybernetics society, pp. 199-204, 2005.

[5] R. Malarvizhi, A. Thanamani," K-NN Classifier Performs Better Than K-Means Clustering in Missing Value Imputation", IOSR Journal of Computer Engineering (IOSRJCE), vol. 6, pp. 12-15, Nov. - Dec. 2012.

[6] N. Devi, Balamurugan.S, Swathi U.V, "An amalgam KNN to predict Diabetes Mellitus", in proc. ICE-CCN, 2013, pp. 691-695.

[7] A. Kusiak, Member, IEEE, "Feature Transformation Methods in Data Mining", IEEE transactions on electronics packaging manufacturing, vol. 24, no. 3, July 2001.

[8] Lukasz A. Kurgan, Member, IEEE, and Krzysztof J. Cios, Senior Member, IEEE, " CAIM Discretization Algorithm", IEEE transactions on knowledge and data engineering, vol. 16 (2), February 2004.

[9] Cheng-Jung Tsai, Chien-I. Lee, Wei-Pang Yang, "A discretization algorithm based on Class-Attribute Contingency Coefficient", vol. 178, pp. 714-731, 1 February 2008.

[10] Wenyu Qu, Deqian Yan, Yu Sang, Hongxia Liang, Masaru Kitsuregawa, and Keqiu Li, "A Novel Chi2 Algorithm for Discretization of Continuous Attributes", Vol. 4976, pp. 560-571, 2008.

[11] J. Ge, Y. Xia and Y. Tu, "A Discretization Algorithm for Uncertain Data", DEXA'10 Proceedings of the 21st international conference on Database and expert systems applications Part II, Springer 2010 pp. 485-499.

[12] Qiusha Zhu, Lin Lin, Mei-Ling ShyuShu-Ching Chen "Effective Supervised Discretization for Classification based on Correlation Maximization", in proc. IEEE, 3-5 Aug.2011, pp. 390-395.

[13] S. Zhou, S. Zhang and G. Karypis "An Enhanced Class-Attribute Interdependence Maximization Discretization", Springer-Verlag Berlin Heidelberg, pp. 465–476, 2012.

[14] Shivani V. Vora and R.G.Mehta,"MCAIM: Modified CAIM Discretization Algorithm for Classification", International Journal of Applied Information Systems", vol. 3, 3(5), pp. 42-50, July 2012.

[15] Trie Website: http://en.wikipedia.org/wiki/Trie.

[16] R. de la Briandais, "File searching using variable length keys", in Proc. Western Joint Computer Conference, AFIPS Press, Montvale, NJ, USA, vol. 15, 1959.

[17] V. Aho, J. E. Hopcroft and J. D. Ullman, Data Structures and Algorithms, Addison Wesley, Reading, Massachusetts, 1983.

[18] J. I. Aoe, K. Morimoto, and T. Sato, An efficient implementation of Trie structures, SOFIWARE—PRACTICE AND EXPERIENCE, vol. 22(9), pp. 695–721, Sept. 1992.

[19] D. E. Knuth, the Art of Computer Programming, 2nd ed., vol. 3, Addison-Wesley, Massachusetts, 1973.

[20] T. C. Bell, J. G. Cleary, and I. H. Witten, Text Compression, Prentice-Hall, Englewood Clis, New Jersey, 1990.

[21] R. A. Baeza-Yates and G. Gonnet, "Fast text searching for regular expressions or automaton searching on Tries", Journal of the ACM, vol. 43(6), pp. 915-936, 1996.

[22] J. L. Peterson, "Computer programs for detecting and correcting spelling errors", Communications of the ACM, pp. 676-686, Dec. 1980.

[23] V. Aho, R. Sethi and J. D. Ullman, Compilers Principle Techniques and Tools, Addison Wesley, Reading, Massachusetts, 1986.

[24] P. Flajolet and C. Puech, "Partial match retrieval of multidimensional data", Journal of the ACM, vol. 33(2), pp. 371-407, 1986.

[25] R. L. Rivest, "Partial match retrieval algorithms", SIAM Journal of Computing, vol. 5(1), pp. 19-50, 1976.

[26] S. Nilsson and G. Karlsson, "IP-address lookup using LC-Tries", IEEE Journal on Selected Areas in Communication, vol. 17(6), pp. 1083-1092, June. 1999.

[27] Zailani Abdullah, Tutut Herawan, and Mustafa Mat Deris, Visualizing the Construction of Incremental Disorder Trie Itemset Data Structure (DOSTrieIT) for Frequent Pattern Tree (FP-Tree). Visual Informatics: Sustaining Research and Innovations Lecture Notes in Computer Science, Springer, vol. 7066, pp. 183-195, 2011.

[28] Noraziah, Zailani Abdullah, Tutut Herawan and Mustafa Mat Deris. Scalable Technique to Discover Items Support from Trie Data Structure. Information Computing and Applications Lecture Notes in Computer Science, Springer, vol. 7473, pp. 500-507, 2012.

[29] Professor Andrew W. Moore, "Decision Trees", school of computer science Carnegie Mellon University.

[30] Xindong Wu, Vipin Kumar et al., "Top 10 algorithms in Data mining", knowl inf Syst, pp. 1-37, 2008.

[31] Chen Jin Luo, De-lin, Mu Fen-xiang, "An Improved ID3 Decision Tree algorithm",4th International Conference on Computer Science & Education,2009, pp. 127-130.

[32] Miao Wang, Ruimin Chai, "Improved Classification Attribute Selection Scheme for Decision Tree", Computer Engineering and Application, pp. 127-129, 2010.

[33] Rong Wang, "New Decision Tree Algorithm", Science Technology and Engineering, vol. 9, 2009.

[34] Quinlan J R,"C4.5 program for machine learning", San Marteo Morgan Kaufmann Publishers, pp.21-30, 1993.

[35] C5.0 website: http://rulequest.com/see5-info.html

[36] Anand Bahelty, "extension and evaluation of Id3-decision tree algorithm", dept. of computer Science, University of Maryland college park.

[37] Schlimmer, J. C., & Fisher, D. "A case study of incremental concept induction", Fifth National Conference on Artificial Intelligence", 1986, pp. 496-501.

[38] Utgoff, P.,"ID5: An incremental ID3", Fifth International Conference on Machine Learning, 1988, pp. 107-120.

[39] Utgoff, P.E.,"Incremental induction of decision trees", Machine Learning 4, pp. 161-186, 1989.

[40] Domingos, P., Hulten, G.,"mining high-speed data streams", KDD, ACM press, New York, USA, pp. 71-80, 2000.

[41] Hulten, G., Spencer and L. Domingos, P. "Mining time-changing data streams", ACM press, New York, pp. 97-106, 2001.

[42] Gamma, J., Fernandes, R., and Rocha,"Decision trees for mining data streams ", Intelligent Data Analysis, pp. 23-45, 2006.

[43] M R lad et al.,"A Novel Tree based Classification", International Journal of Engineering Science & Advanced technology", vol. 2(3) pp.581-586, may-Jun 2012.

Discretization of Temporal Data: A Survey

P. Chaudhari[*1], D. P. Rana[*2], R. G. Mehta[*3], N. J. Mistry[*], M. M. Raghuwanshi[#]

[*]*SVNIT, Surat, India*
[1]pratikshachaudhari29@gmail.com
[2]dpr@coed.svnit.ac.in
[3]rgm@coed.svnit.ac.in

[#]*RGCET, Nagpur, India*

Abstract—**In real world, the huge amount of temporal data is to be processed in many application areas such as scientific, financial, network monitoring, sensor data analysis. Data mining techniques are primarily oriented to handle discrete features. In the case of temporal data the time plays an important role on the characteristics of data. To consider this effect, the data discretization techniques have to consider the time while processing to resolve the issue by finding the intervals of data which are more concise and precise with respect to time. Here, this research is reviewing different data discretization techniques used in temporal data applications according to the inclusion or exclusion of: class label, temporal order of the data and handling of stream data to open the research direction for temporal data discretization to improve the performance of data mining technique.**

Keywords— **Temporal data, Discretization, Supervised, Incremental, Nonparametric**

I. INTRODUCTION

The huge amount of data is to be processed in many application areas such as scientific, financial, network monitoring, sensor data analysis [1], [2]. Data mining is important to analyse huge amount of data and with time oriented data, data analysis can be done better and more naturally. Recently, the increasing usages of temporal data in various applications have got focus on research in the field of data mining. Temporal data can be easily obtained and available from scientific and financial applications like ECG, diabetes, daily weather data, sales information and stocks information. A time series is a collection of observations made chronologically. The characteristics of time series data includes: huge in size, high dimensions and updated continuously.

Moreover temporal data have characteristics of numeric and continuous nature, which has to be always considered as a group instead of individual field. There are number of research is going on temporal data, to find similarity, to segment and to search, to reduce in dimension. Temporal discretization refers to the discretization of time series, as a preprocessing step in transforming the temporal data into timely intervals. An effective discretization method not only reduces the dimensionality of data and improve the efficiency of data mining and machine learning algorithm, but also make the knowledge extracted from the discretized dataset more compact, easy to understand and useful. Many data mining algorithms can benefit from a discrete representation of the original data set as the numbers of values for a feature are many, but after the discretization, the number of discrete values is less compare to original number of values [1], [2].

Discretization has a number of advantages: Discrete features reduce memory usage and thus increase representation of the knowledge as data is simplified to understand and with this application of mining technique or knowledge retrieval methods become faster and perfect [2].

There are several ways by which discretization methods can be classified: Splitting versus Merging, Global versus Local, Supervised versus Unsupervised, Static versus Dynamic, and non-Incremental versus Incremental [2], [3], [4], [5].

- The Splitting methods, is a top-down approach of discretization in which one start with an empty set of cut points and gradually divide the interval and subintervals to obtain the discretization. In contrast, the merging method is a bottom-up approach in which it considers all possible cut points and then eliminate these cut points by merging intervals.
- Local methods produce partitions that are applied to localized regions of instance space. Global methods, on the other hand, use the entire instance space and form a mesh over the entire n-dimensional continuous instance space, where each feature is partitioned into regions independent of other attributes.
- Unsupervised methods carry out discretization without the knowledge of class label, whereas the supervised methods utilize the class information to carry out the discretization.
- Static discretization methods require some parameter indicating the maximum number of desired intervals to discretize a feature. Dynamic methods conduct a search through the space of possible k values for all features simultaneously.
- Non-incremental methods consider only the available historical data values. But, with time, the data values that can be completely new are not considered. While in incremental methods they try to cover the new values also.

Here, this research is discussing the different data discretization methods used for different temporal data applications that may have a) class label or no class labels, b) nature of continuous data or static data and c) considering the temporal order or not considering the temporal order.

The organization of rest of this research is as follows: Section 2 discusses the literature review of temporal static data discretization methods which is not considering the temporal order of the timely data and the other discretization methods which is considering the temporal order of the data

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

under the unsupervised and supervised discretization category followed by the comparative analysis of these discretization methods for defined parameters. Final Section 3 summarizes the research for the future scope.

II. DISCRETIZATION

Many data mining algorithms and tasks can benefit from a discrete representation of the original data set. Discrete representation is more comprehensive to human and can simplify, reduce computational costs and improve accuracy of many algorithms. Discretization is the process of transforming continuous space valued series X={x1,x2,...,xn} into a discrete valued series Y={y1,y2,...,yn}. Discretization can be performed recursively on an attribute. The main part of the discretization process is choosing the best *cut points* which split the continuous value range into discrete number of bins usually referred to as *states*.

From the literature review, it is found that most of discretization methods which are used can be majorly categorized as unsupervised and supervised discretization. Here, this section is discussing discretization methods according to this category.

A. Unsupervised Discretization

When in dataset class information is not available for time series, unsupervised methods are needed. Two common methods used in most of the applications are Equal Width Discretization (EWD) and Equal Frequency Discretization (EFD) [1]. Other than these are K-means Clustering [1], [3] SAX [6], Frequency Dynamic Interval Class (FDIC) [7], where methods: EWD, EFD and K-means clustering are static methods which are not considering temporal order of data.

1) Equal Width Discretization (EWD): EWD is a simplest discretization method that divides the range of observed values for a feature into k equal sized bins, where k is a parameter provided by the user [4]. The process involves finding values as the minimum (Vmin) and maximum (Vmax). The interval is computed by dividing the range of observed values for the variable into k number of equally sized bins using the formula Interval=(Vmax-Vmin)/k, where k is a parameter supplied by the user and Boundaries = Vmin+(i * interval) for the i = 1...k-1 boundaries [3]. However, this method of discretization is sensitive to outliers that may drastically skew the range.

The limitations of this static, unsupervised and very simple method are: a) It is the parametric method as the required number of interval is needed from the user. b) It follows the characteristics of how data values are distributed as some intervals may contain much more data points than other and can produce overlapping.

2) Equal Frequency Discretization: The equal-frequency algorithm determines the minimum and maximum values of the discretized attribute of n values, it sorts all values in increasing frequency order and then divides the sorted data values into *k* bins such that each interval approximately contains *n/k* data values with adjacent internal values. For equal frequency, many occurrences of a continuous value could cause the occurrences to be assigned into different bins that cause the problem of overlapping.

This method is advantageous and tries to overcome the limitations of the equal-width interval discretization by dividing the domain in intervals with the same distribution of data points. The problems with this static, unsupervised and parametric method is that it is not always possible to generate exactly k equal frequency intervals because it tries to place the data instance with identical value in the same interval.

The author Chaves has applied this method for image histogram discretization [8].

3) K-means Clustering: The K-means clustering method is widely used for temporal data discretization in number of applications as it helps to find natural groups and one of the data mining techniques [2], [3], [5]. The author Salvador used the K-means clustering to identify the number of states in a time series dynamically [9], the author Z. Liang et al. used this method to partition the values of the attributes like temperature, salinity, pH, etc. to detect the correlation between environmental factors and ecological events [10].

In K-means clustering, a Euclidean distance measure is a simple distance function used to cluster the data into *k* clusters which are represented by the centroids. The clustering algorithm begins with a random or more educated choice (more efficient due to the sensitivity of the clustering process to the initial selection) of clusters centroids. The second step is to assign each data point to the cluster that has the closest centroid. After every data point has been assigned, the *k* centroids are recalculated as the mean value of each cluster. Two steps are repeated until no data point is reassigned or the *k* centroids no longer change. The resulting clusters or centroids are used as the states of the discretization process.

This static method is advantageous than EWD and EFD methods as the data are grouped together as per their characteristics. But this can be the limitations also and other than this are: a) It is the parametric method, as the required number of interval value k is to be provided by the user. b) The number of intervals of discretization is dependent upon the given k value and seed value of cluster.

4) Symbolic Aggregate Approximation (SAX): SAX is the advanced method especially designed for temporal data. The author Keogh et al. applied SAX discretization for improvement of pattern finding performance [11].

The SAX representation of a time series is given by the author Lin et al. in 2003 [6] to convert the piecewise aggregate approximation to symbols. The authors has achieved this by dividing the vertical y-axis into equi probable parts, which replaced by symbol. To do so, parameters like subsequence length and the number of symbols. SAX helps to reduce the dimensionality and lower bounds the distance between any two vectors in the SAX representation is smaller than, or equal to, the distance between these two vectors in the original space.

Before discretization, the author has removed distortions by normalizing each time series and that can follow the Gaussian

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

distribution. Then for the given number of segmentation, transform time series into partial aggregate approximation. And now breakpoints are decided so that they produce equal-sized areas under the Gaussian curve using the breakpoint lookup table that helps to divide the amplitude values of the time series into required number of equi-probable regions. Then depending upon the breakpoint lying area symbols are provided to generate the complete symbol string.
The advantages of this method are a) It is reducing dimensionality depending upon alphabet size. b) It is lower bounding the distance between any two vectors. The SAX method is suitable for most classic data mining tasks like classification and clustering.

The extended-SAX is proposed by the same author in 2007 [12] which included the min and max value into the consideration. Modification to SAX method, FAST-SAX is proposed by M. Fuad where each series in the database is represented by a first-degree polynomial, which is the approximating function for all the time series in the databases [13]. The distances between the time series and their approximating function are computed and stored instead of time series to perform faster than original SAX method.

5) Frequency Dynamic Interval Class (FDIC): FDIC is advanced and novel approach of data discretization method which is dynamic in nature presented by A. M. Ahmed et al. [7]. The method consists of two phases. First is dynamic interval class and second is interval merging phase. A dynamic interval phase which tries to generate a number of intervals from data information itself, then generated width intervals by considering the intervals of frequency distribution of the pattern and distance of interval patterns and uses automatically computed threshold based on means distribution to determine the number and the length of intervals.

The second phase merging phase which begins with calculation of distance between known interval class and unknown class. In this phase unknown class intervals that have less distribution, are merged with the nearest point with satisfied minimum threshold. The SV-kNNC algorithm is used to determine the new cut point of unknown class interval. It is dependent on nearest known intervals which have more distribution on data. As a result, when discretization ends, numerical continuous attribute values are transformed into discrete ones based on the dataset characteristics.

This nonparametric and bottom up, method is more advantageous over other methods as the data are grouped together as per their characteristics and frequency of values. But only limitation of this method is it is not incremental.

B. Supervised Discretization

There are numbers of supervised discretization methods based on entropy interested readers can check with [2], [3], [14] but here the methods based on clustering [15] and based on Class-Attribute Contingency Coefficient [16] are discussed as they are according to temporal data characteristics and also with class information.

1) Clustering based Discretization: This method is taking the advantage of class label and natural grouping together. It is based on clustering, that discretizes the data with the knowledge of both classes and clusters [15]. In this method two clustering algorithms used – the K-means clustering approach with Euclidean distance metric as the similarity measure and the shared nearest neighbour (SNN) clustering algorithm. In the K-means clustering approach the number of clusters is kept approximately equal to the number of classes. Each data instance is assigned to a particular cluster and this is named as the 'pseudo-class'.

Clustering provides the intrinsic grouping of the unlabeled data. Thus, the cluster id captures the interdependencies in the data. So, there are two class features, one provided with data, C, and the other, pseudo-class C'. Then entropy based discretization applied to search for the partition of the value range of a continuous feature so as to minimize the uncertainty of the class variable conditioned on the discretized feature variable. This method results in two intervals and is applied recursively to each subsequent subinterval until the stopping criterion is met.

This method is advantageous because in discretization of continuous variables simultaneously using the class information and cluster based 'pseudo-class' information generally better than that based on the class information alone.

2) Class Attribute Contingency Coefficient Discretization (CACC) Method: This is supervised and top-down discretization method based on Class-Attribute Contingency Coefficient by Lee et al. [14]. It calculates CACC value which is used in discretization of continuous data to measure the interdependence between variables.

For each attribute, CACC first finds the maximum and minimum of attributes and then forms a set of all values of each attribute in the ascending order. For all possible interval boundaries and all the midpoints of all the adjacent boundaries in the set are obtained and keeping the maximum CACC value and then partition this attribute accordingly into intervals. The CACC discretization method raise the quality of the generated discretization scheme by extending the idea of contingency coefficient and give better result as considers different attributes together than the consideration of individual attribute, which is the main characteristic of temporal data. But, major limitation is that it is not considering temporal order of data.

From the study of various discretization methods, we derived the following analysis as described in Table 1 and we found that enough research work is done in the area of unsupervised discretization, but very less work is done in the supervised discretization together with incremental approach.

III. Summary and Future Scope

Discretization of data plays an important role in data preprocessing before applying a number of data mining algorithms on the real valued data sets. Here, briefly introduced the need of discretization with the idea and drawbacks of some methods under supervised or unsupervised

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

category for the temporal data. Analysis has been given based on different issues of discretization which shows that still research area is open to consider some issues like nonparametric automatic discretization approach to consider continuous streaming data with the consideration of temporal order.

TABLE I ANALYSIS OF TEMPORAL DATA DISCRETIZATION METHOD

Category	**UnSupervised**		**Supervised**		
Methods / Criteria	**EWD, EFD, K-Means**	**SAX, Extended-SAX, Fast-SAX**	**FDIC**	**CACC**	**Clustering based discretization**
Supervised or Not	No	No	No	Yes	Yes
Consideration of Temporal order	No	Yes	No	No	No
Top Down / Bottom Up	Top Down	Top Down	Bottom Up	Top Down	Top Down
Parametric	Yes	Yes	No	No	Yes
Overlapping solved	No	No	Yes	Yes	No
Incremental	No	No	No	Yes	No

REFERENCES

[1] Han and Kamber, "Data Mining: Concepts and Techniques", Second Edition.

[2] T. Fu, "A review on time series data mining", Engineering Applications of Artificial Intelligence, pp. 164–181, 2010.

[3] R. Dash, R. L. Paramguru and R. Dash, "Comparative Analysis of Supervised and Unsupervised Discretization Techniques", International Journal of Advances in Science and Technology, Vol. 2, No. 3, 2011.

[4] J. Dougherty, R. Kohavi and M. Sahami, "Supervised and unsupervised discretization of continuous features", International Conference on Machine Learning, pages 194-202, (1995).

[5] H. Liu, F. Hussain and C. L. Tan, "Discretization: An Enabling Technique", Data Mining and Knowledge Discovery, Vol. 6, pp. 393–423, 2002.

[6] J. Lin, E. Keogh, S. Lonardi and B. Chiu, "A symbolic representation of time series, with implications for streaming algorithms", in Proc. of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, San Diego, California, pp. 2–11, June 2003.

[7] A. M. Ahmed, A. A. Bakar and A. R. Hamdan, "Dynamic data discretization technique based on frequency and K-Nearest Neighbour algorithm", in Proceedings of the 2nd Conference on Data Mining and Optimization, Malaysia, pp. 27–28, 2009.

[8] R. Chaves, J. Ramirez and J.M. Gorriz, "Integrating discretization and association rule-based classification for Alzheimer's disease diagnosis", Expert Systems with Applications, pp. 1571–1578, 2013.

[9] S. Salvador, P. Chan and J. Brodie, "Learning States and Rules for Time Series Anomaly Detection", in Proc. of the 17th Intl. Flairs Conference, 2004.

[10] Z. Liang, T. Xinming and J. Wenliang, "Temporal Association Rule Mining Based On T-Apriori Algorithm and its Typical Application", International Symposium on Spatial-Temporal Modeling Analysis, Vol. 5 No. 2, 2005.

[11] E. Keogh, S. Lonardi, Y. C. Chiu, Finding surprising patterns in a time series database in linear time and space, in Proc. of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and DataMining, pp. 550–556, 2002.

[12] J. Lin, E.J. Keogh, L. Wei and S. Lonardi, "Experiencing sax: a novel symbolic representation of time series", Data Mining Knowledge Discovery, Vol. 15 (2), pp. 107–144, 2007.

[13] M. Fuad and P. Marteau, "Towards A Faster Symbolic Aggregate Approximation Method", Journal CoRR, Vol. abs/1301.5871, 2013.

[14] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations", in Proceedings of the Fifth Symposium on Math, Statistics, and Probability, Berkeley, CA: University of California Press, pp. 281–297, 1967.

[15] A. Gupta, K. G. Mehrotra and C. Mohan, "A Clustering-based discretization for supervised learning", Statistics and Probability, pp. 816–824, 2010.

[16] C. Tsai, C. Lee and W. Yang, "A discretization algorithm based on Class-Attribute Contingency Coefficient", Information Sciences, Vol. 2, pp. 714–731, Sept 2008.

AUTHORS PROFILE

P. Chaudhari is M.Tech scholar in Computer Engineering Department at SVNIT, Surat, Gujarat, India.

D. P. Rana is Assistant Professor at Computer Engineering Department, S. V. National Institute of Technology, Surat, Gujarat, India and is currently pursuing her PhD degree. Her research interest is in the field of security in web applications, computer architecture, database management system, data mining and web data mining. She is a life member of ISTE and CSI.

R. G. Mehta is Associate Professor at Computer Engineering Department, S. V. National Institute of Technology, Surat, Gujarat, India and is currently pursuing her PhD degree. Her research interest is in the field of security in web applications, computer architecture, database management system, data mining and web data mining. She is a life member of ISTE and CSI.

N. J. Mistry is Professor at Civil Engineering Department, S. V. National Institute of Technology, Surat, Gujarat-395007, India. He is a member of CES.

M. M. Raghuwanshi is working as a principal at Rajiv Gandhi College of Engineering and Research, Nagpur, India. He completed his PhD in Computer Science, 2007, at VNIT, Nagpur, India. He is a member of IEEE, ISTE and CSI.

IJCSIS AUTHORS' & REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore,Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGLIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Mrs Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Mr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Mr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Mr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Mr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Mr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Mr. P. Vasant, University Technology Petronas, Malaysia
Mr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Mr. Praveen Ranjan Srivastava, BITS PILANI, India
Mr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Mr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktesh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia
Mr. Tirthankar Gayen, IIT Kharagpur, India
Ms. Huei-Ru Tseng, National Chiao Tung University, Taiwan

CALL FOR PAPERS
International Journal of Computer Science and Information Security

IJCSIS 2014
ISSN: 1947-5500
http://sites.google.com/site/ijcsis/

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc,), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on
its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embeded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://sites.google.com/site/ijcsis/authors-notes .

ISSN 1947 5500
http://sites.google.com/site/ijcsis/